PROBABILITY AND ITS APPLICATIONS TO ANTISUBMARINE WARFARE

Billy Joe Smith

NAVAL POSTGRADUATE SCHOOL
Monterey, California

GROUP PROJECT

PROBABILITY AND ITS APPLICATIONS
TO ANTISUBMARINE WARFARE

by

Billy Joe Smith

and

Larry Wayne Vice

March 1975

Thesis Advisor: Carroll O. Wilde

Approved for public release; distribution unlimited.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Probability and Its Applications to Antisubmarine Warfare		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis March 1975
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Billy Joe Smith and Larry Wayne Vice		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE March 1975
		13. NUMBER OF PAGES 98
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Probability
Antisubmarine Warfare
ASW
Operational Probability Applications
ASW Applications

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A basic course with applications of probability to ASW, the course consists of six lesson plans and a "Study Guide." The lesson plans are designed to give the/an instructor guidance in what to teach, the depth required and objectives the student should be able to accomplish. The "Study Guide" provided is for the use of both the instructor and the student, and it should serve as a basic text for the course. Current

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
(Page 1) S/N 0102-014-6601 |

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20.
Abstract continued--

ASW tactical publications were examined by the authors while developing the course, and as many of the probability applications and as much associated probability terminology from these sources as practicable (and when this could be accomplished at the "unclassified" level) are incorporated in the course.

Probability and Its Applications to Antisubmarine Warfare

by

Billy Joe Smith
Lieutenant, United States Navy
B. S., University of Oklahoma, 1966

and

Larry Wayne Vice
Lieutenant, United States Navy
B. S., University of Oklahoma, 1968

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS TECHNOLOGY

from the

NAVAL POSTGRADUATE SCHOOL
March 1975

ABSTRACT

A basic course with applications of probability to ASW, the course consists of six lesson plans and a "Study Guide." The lesson plans are designed to give the/an instructor guidance in what to teach, the depth required and objectives the student should be able to accomplish. The "Study Guide" provided is for the use of both the instructor and the student, and it should serve as a basic text for the course. Current ASW tactical publications were examined by the authors while developing the course, and as many of the probability applications and as much associated probability terminology from these sources as practicable (and when this could be accomplished at the "unclassified" level) are incorporated in the course.

ACKNOWLEDGEMENT

The authors would like to express their indebtedness to their thesis advisor, Professor Carroll O. Wilde, for his many ASW examples and problems, and his continued support, guidance, enthusiasm and encouragement throughout this thesis effort.

We thank our wonderful wives for their patience during this endeavor.

PROBABILITY AND ITS APPLICATION TO ANTISUBMARINE WARFARE

Introduction

One of the most important characteristics of modern ASW (Antisubmarine Warfare) is the existence of so many uncertainties and unknowns. Examples of these "unknowns" abound in almost every aspect of this warfare area. To a Tactical Coordinator on a search aircraft the "unknown" may be the evasive maneuvers an enemy submarine skipper will employ, how many of the sonobuoys he has aboard his aircraft are defective, or which portion of a search region offers the best prospects for a contact. To a CIC officer the presence of a submarine within detection range may be unknown because of excessive ambient noise; at the time a contact is gained it may be unknown to him whether the contact is in fact a submarine, and then the target location accuracy may be unknown because of uncertainties in underwater sound propagation.

Because of the existence of so many unknowns, many Naval Officers who work in ASW are tasked with making action recommendations and action decisions, which are based on predictions of the unknown in the light of all that <u>is</u> known. Therefore, success in the overall ASW effort is directly related to the capability within the ASW-related officer corps in two key areas: (i) making predictions concerning the unknown, and (ii) understanding the concomitant degree of accuracy and degree of uncertainty of such predictions, so that the best possible decisions may then be made.

Probability theory provides a logical and scientific basis for prediction and decision making. It is not surprising, therefore, that

probabilistic concepts and the language of probability are in widespread everyday usage in the ASW community today. Even the basic references used in ASW, such as ComTac publications, ASW handbooks, Tactical Reference manuals, TacAids, ASRAPS and AESPAWS, use the language and concepts of probability theory freely.

Because of the widespread use of probability in the tasks assigned to the ASW-related Officer, his effectiveness should be enhanced by an increased understanding of probability theory and a better appreciation of its benefits in ASW. This report constitutes a "tutorial paper" designed to provide a means by which ASW-related officers may better understand application of probability in modern ASW. The objective of this project has been to distill from the general field of probability those elements that are vital in ASW applications, and to describe these elements, along with typical examples of their application, in an organized and readable presentation that is accessible to any ASW-related Officer, with the only prerequisites being basic algebra and geometry as applied to tactical situations in his working environment.

This paper has been organized into the format of an academic course, including lesson plans, study guides and suggested textbooks. This organization was adopted because it is expected that this paper will serve as the basis for a future short course that will be available to the Fleet through the Continuing Education Office at the Naval Postgraduate School. Because of the possible misunderstanding that may result from the "course format" adopted here, it is emphasized that <u>the content of this paper does not constitute a Naval Postgraduate School course</u>. It is intended that an official course be developed from these notes; this, however, involves an evolutionary process, including testing of the materials with actual students, and subsequent revisions.

The Lesson Plans provide the objectives of each unit; they indicate what is to be learned, the depth required, and the behavioral goals. The Study Guides serve as the basic textbook. In addition, Schaum's Outline of Theory and Problems of Probability by S. Lipschutz (1965) and Schaum's Outline of Theory and Problems of Statistics by M. R. Spiegel (1961), both by the McGraw-Hill Book Company, are very helpful; these texts are inexpensive and easy to follow, as they contain many samples with detailed solutions. Reference is made to the appropriate sections in the Schaum's texts at the end of each chapter.

The primary material presented is the probability theory required for successful ASW application, and examples to illustrate the application. The probability theory is preceded by a presentation of sufficient statistical procedures to serve as "first principles" on which the probability is based.

Lesson Plan No. S-V-1

Time: 65 min.

Date:_______________

Instructor:_________________________

Subject Area: Probability and Its Applications to Anti-submarine Warfare

Topic: Data Reduction, Frequency Distribution

Reference: ASW Study Guide; Schaum's Outline Series, <u>Theory and Problems of Statistics</u>

Title: "Sample Data Reduction and Frequency Distributions"

Materials: Transparencies of Table 1, Table 2, Table 3, Figure 1, Figure 2, and Figure 3

Objectives:

(1) Acquaint the student with the concepts of statistical <u>population</u> and <u>samples</u> from the population; the student should be able to explain these concepts by means of simple examples.

(2) Discuss the basic terms required to describe frequency distributions, and show mathematically how they are utilized. The student should be able to determine the frequency distribution from a given set of raw data.

Introduction: Instructor should inform students of purpose of lesson and briefly highlight important points to be covered.

Presentation: Discuss the following points:

1. Population
2. Sample
3. Methods of describing a sample and estimating characteristics of the population
4. Raw data
5. Arrays
6. Classes or categories
7. Class frequency and frequency table
8. Class intervals and class limits
9. Class boundaries
10. Size or width of a class interval

11. Class mark

12. General rules for forming frequency distributions:

    a. The range

    b. Number of class intervals

    c. Number of observations falling into each class interval

13. Histograms and frequency polygons

14. Relative frequency distributions

15. Cumulative frequency distributions or ogives

16. Relative cumulative frequency distributions

17. Types of frequency curves

    a. Symmetrical or bell-shaped

    b. Skewed to the right

    c. Skewed to the left

    d. U-shaped

    e. Bimodal

    f. Multimodal

Summary: Review material covered, reemphasizing key points. Ask questions to determine effectiveness of instruction.

Lesson Plan No. S-V-2

Time: 115 min.

Date:________________

Instructor:________________________

Subject Area: Probability and Its Applications to Antisubmarine Warfare

Topic: Measures of Central Tendency and Dispersion

Reference: ASW Study Guide; Schaum's Outline Series, Theory and Problems of Statistics

Title: "Special Measures of Sample Data--Central Tendency and Dispersion"

Materials: Transparency of Figures 4 through 6

Objectives: (1) Define central tendency and its measurement from sample data. Illustrate and develop the sample mean, median, and mode. The student shall be able to calculate the mean, median, and mode from given sample data.

(2) Define dispersion and its measurement from sample data. Illustrate, and develop the listed measures of dispersion so that the student is capable of calculating the range, standard deviation and variance from given sample data.

Introduction: Instructor should inform students of purpose of lesson and briefly highlight important points to be covered.

Presentation: Discuss the following points:

1. Index or subscript notation

2. Summation notation

3. Measures of central tendency

   a. Mean

      (1) (Arithmetic) Mean
      (2) Weighted (arithmetic) mean

   b. Median

   c. Mode

4. The concept of dispersion of data

5. Specific measures of dispersion

   a. Range

   b. Standard Deviation

(1) Formulation of
(2) Short method of computing

c. Variance

d. Sigma-values for Normal probability distributions with corresponding probability values.

Summary: Review material covered, reemphasizing key points. Ask questions to determine effectiveness of instruction.

Lesson Plan No. S-V-3

Time: 20 min.

Date:________________

Instructor:__________________________

Subject Area: Probability and Its Applications to Anti-submarine Warfare

Topic: Combinatorial Analysis

Reference: ASW Study Guide; Schaum's Outline Series, Theory and Problems of Statistics

Title: "Combinatorial Analysis"

Objective: Develop the concept of COMBINATORIAL ANALYSIS by defining and illustrating, with a simple example, the basic elements of this analysis. The student should be able to perform combinatorial analysis on a given example by utilizing the fundamental counting principle, and calculating permutations/combinations to find the number of possible outcomes for the example.

Introduction: Instructor should inform students of purpose of lesson and briefly highlight important points to be covered.

Presentation: Discuss the following points:

1. The fundamental counting principle
2. Factorial n
3. Combinations
4. Permutations

Summary: Review material covered, reemphasizing key points. Ask questions to determine effectiveness of instruction.

Lesson Plan No. S-V-4

Time: 150 min.

Date:________________

Instructor:________________________

Subject Area: Probability and Its Applications to Anti-submarine Warfare

Topic: Elementary Probability Theory

Reference: ASW Study Guide; Schaum's Outline Series, Theory and Problems of Statistics; and Schaum's Outline Series, Theory and Problems of Probability

Title: "Elementary Probability Theory"

Materials: Transparencies of Figures 5 through 12

Objectives: (1) Develop student intuition in the basic concepts of probability, i.e. his insight into what probability is. Develop the concepts of basic probability through the use of the proportional areas notion of probability. The student should be able to compute simple probabilities given sample data and simple statement of necessary conditions.

(2) Develop the concept of CONDITIONAL PROBABILITY as an improved probability resulting from having previously eliminated part of the overall sample space to be "drawn on," through use of prior information. The student should be able to compute conditional probabilities utilizing any of the three listed methods.

(3) Develop the concept of STATISTICAL INDEPENDENCE, and delineate between this and MUTUALLY EXCLUSIVE EVENTS. Also, show the utility of statistical independence of events in computing a simple joint probability. Given two events, the student should be able to state whether or not they are statistically independent.

Introduction: The instructor should inform students of purpose of lesson and briefly highlight important points to be covered.

Presentation: Discuss the following points:

1. Elementary Probability Theory

a. Background and concepts of probability

(1) "Odds" or ratio concept

(2) Proportional areas concept

b. Sample Space

c. Probability Experiment

d. Outcomes

e. Events

f. Probability Function

g. Probability Complement

2. Conditional Probability

a. Concept of

b. Methods of computing

(1) By definition

(2) "Stating and Solving"--"Simply writing it down"

(3) Bayes' Law

3. Statistical Independence

a. Independent Events; Probability of Independent events

b. Independent Events vs. Mutually Exclusive Events

c. Implications/applications of Statistical Independence

Summary: Review material covered, reemphasizing key points. Ask questions to determine effectiveness of instruction.

Lesson Plan No. S-V-5

Time: 100 min.

Date:________________

Instructor:______________________

Subject Area: Probability and Its Applications to Anti-submarine Warfare

Topic: Random Variables

Reference: ASW Study Guide; Schaum's Outline Series, Theory and Problems of Probability

Title: "Random Variables"

Materials: Transparencies of Table 5, Table 6 and Figures 15 through 19

Objectives: Introduce the concept of random variables, and develop random variables as a device for assigning names to the outcomes of a probability experiment. The student should be able to demonstrate how a random variable is used by means of examples.

Introduction: Instructor should inform students of purpose of lesson and briefly highlight important points to be covered.

Presentation: Discuss the following points:

1. Random variable
   a. discrete
   b. continuous
   c. density function
   d. cumulative distribution function (CDF)
   e. mean
   f. variance
   g. standard deviation

Summary: Review material covered, reemphasizing key points. Ask questions to determine effectiveness of instruction.

Lesson Plan No. S-V-6

Instructor:______________________

Time: 550 min.

Subject Area: Probability and Its Applications to Anti-submarine Warfare

Date:________________

Topic: Basic Statistical Distributions

Reference: ASW Study Guide; Schaum's Outline Series, Theory and Problems of Statistics; and Schaum's Outline Series, Theory and Problems of Probability

Title: "Basic Statistical Distributions"

Materials: Transparencies of Table 7 and Figures 20 through 29

Objectives:

(1) Present examples of some useful statistical distributions.

(2) The student should be able to give a simple description of each of the distributions presented, and compute probabilities for Bernoulli, Binomial, Uniform, Chi-Square, Exponential, Poisson and Normal examples (with the exception of the Bivariate Normal and Time-Varying Normal, which are presented only for student appreciation of "state-of-the-art" applications).

(3) It is intended that this lesson serve to extend the student's understanding of probability functions, random variables, and density functions and cumulative density functions.

Introduction: Instructor should inform students of purpose of lesson and briefly highlight important points to be covered.

Presentation: Discuss the following points:

1. Background on the basics of specifying a statistical distribution, drawing the majority of this information directly from the previous lesson plan.
2. Bernoulli Distribution
3. Binomial Distribution
4. Uniform Distribution
5. Normal Distributions
   a. Normal Distribution
   b. Bivariate Normal Distribution

c. Time-Varying Normal Distribution

6. Chi-Square Distribution

7. Exponential Distribution

8. Poisson Distribution

Summary: Review material covered, reemphasizing key points. Ask questions to determine effectiveness of instruction.

STUDY GUIDE

PROBABILITY AND ITS APPLICATIONS
TO ANTISUBMARINE WARFARE

INTRODUCTION

Probability theory originated in the early seventeenth century with investigations of various games of chance. Since then many scholars have made contributions to the theory. Probability theory was finally axiomatized in the twenties and thirties of this century. This axiomatic development, called Modern Probability Theory, provided for a precise formulation of probabilistic concepts, which enabled the theory to be placed on a firm mathematical foundation.

The importance of probability has increased significantly in the past twenty years. Today the notions of probability, and the closely related subject statistics, appear in almost every discipline. Probability is used routinely by many Naval Officers in the line of duty, and plays an important role in Antisubmarine Warfare, especially in the area of search and detection.

This study guide is designed for an introductory course in probability and applications to antisubmarine warfare, with basic algebra as the only prerequisite. The guide may serve as a text for such a course, as lesson plans are also provided, or as a supplement. In addition, since this study guide is complete and self-contained, it could be used for self-study.

Chapters I, II, and III of the study guide cover sample data reduction, frequency distributions, associated measures of central tendency and disperison, and combinatorial analysis. This leads naturally to elementary probability theory, conditional probability, statistical independence and applications, in Chapter IV. Chapter V covers random

variables. Discrete and continuous random variables are covered, with emphasis on the use of tables to determine numerical probabilities. Chapter VI covers the Bernoulli, Binomial, Uniform, Normal, Chi-square, Exponential and Poisson distributions, which are commonly used in antisubmarine warfare applications. Since developing understanding of the applications of these distributions to ASW is the primary "payoff" for this course, Chapter VI is significantly longer and has more applications than the first five chapters.

Each chapter has clear statements of pertinent definitions and principles, together with illustrations and other descriptive material. This is followed by solved problems, which serve to illustrate and amplify the theory, bring into sharp focus those fine points without which the student feels himself on unsafe ground, and provide the repetition of basic principles so vital to effective learning.

CHAPTER I: FREQUENCY DISTRIBUTION

POPULATION AND SAMPLE

In collecting data on characteristics of a group of individuals or objects, such as heights of students in a university, or the number of defective and non-defective sonobuoys produced in a factory on a given day, it is often impossible or impractical to observe the entire group, especially if it is large. Instead of examining the entire group, called the population, one examines a small part of the group called a sample. If a sample is representative of a population, important conclusions about the population can be reached from analysis of the sample. Inferences based on sample data are not absolutely certain because in practice it is usually not known beforehand whether the sample is indeed representative of the population, so the term probability is used in describing conclusions. The term statistics is used in connection with conditions under which such an inference is made.

RAW DATA

Raw data are collected data which have not been organized numerically. An example is the set of heights of 100 students obtained from Naval Academy records. The heights of the 100 students, with repetitions included, are raw data which constitute a sample, and the population from which the sample is taken is the set of heights of the students from the entire listing of the Naval Academy records.

ARRAYS

An array is an arrangement of raw numerical data in ascending or descending order of magnitude.

CLASSES OR CATEGORIES, CLASS FREQUENCY AND FREQUENCY TABLE

In order to comprehend what masses of collected raw data really mean, and to draw conclusions from these data, it is useful to distribute the data into classes or categories. The number of individual data points belonging to a class is called the class frequency. A tabular arrangement of data by classes, together with the corresponding class frequencies, is called a frequency table. Table 1 is the frequency table of the heights (recorded to the nearest inch) of the 100 Naval Academy students.

Height (inches)	Number of Students
(class)	(class frequency)
60 - 62	5
63 - 65	18
66 - 68	42
69 - 71	27
72 - 74	8
	Total 100

Table 1: Heights of 100 Naval Academy Students

CLASS INTERVALS AND CLASS LIMITS

The symbol defining a class, such as 60 - 62 in Table 1, is called a class interval. The end numbers, 60 and 62, are called class limits; the smaller number 60 is the lower class limit and the larger number 62 is the upper class limit. A class interval which has no upper, or no lower, or neither an upper nor a lower class limit is called an open class interval. For example, an open class interval is appropriate for describing torpedoes with a range of "10,000 yards or greater."

CLASS BOUNDARIES

In Table 1 the heights are recorded to the nearest inch, and the class interval 60 - 62 includes all measurements from 59.5000 to 62.5000 inches. These numbers, indicated briefly by the exact numbers 59.5 and 62.5, are called class boundaries; the smaller number 59.5 is the lower class boundary and the larger number 62.5 is the upper class boundary.

SIZE OR WIDTH OF A CLASS INTERVAL

The size or width of a class interval is the difference between the lower and upper class boundaries and is referred to as the class width, class size, or class length. For the data in Table 1, for example, the class width is 62.5 - 59.5 = 65.5 - 62.5 = 74.5 - 71.5 = 3.

CLASS MARK

The class mark is the midpoint of the class interval; it is obtained by adding the lower and upper class limits and dividing by two. Thus the class mark of the interval 60 - 62 is (60 + 62)/2 = 61. The class mark is also called the class midpoint.

GENERAL RULES FOR FORMING FREQUENCY DISTRIBUTIONS

1. Determine the largest and smallest numbers in the raw data and thus find the <u>range</u> (the length of the interval from the smallest to the largest number).
2. Divide the range into a convenient number of class intervals having the same size. The number of class intervals is usually taken between 5 and 20, depending on the data. Whenever possible, class intervals are also chosen so that the class marks or midpoints coincide with actually observed data. However, class boundaries usually do not coincide with actually observed data.
3. Determine the number of observations falling into each class interval; i.e., find the <u>class frequencies</u>. As an example of a frequency distribution, consider the wages in dollars per week paid to enlisted personnel aboard a particular destroyer. Suppose the wages ranged from $50.00 to $120.00 per week with 8 personnel receiving wages between $50.00 and $60.00, 10 between $60.00 and $70.00, 16 between $70.00 and $80.00, 14 between $80.00 and $90.00, 5 between $100.00 and $110.00 and 2 between $110.00 and $120.00. A possible formulation of the corresponding frequency distribution is presented in Table 2 with the range from $50.00 to $119.99, in 7 ten dollar class interval widths and the total number of personnel listed for each class interval.

Wages (dollars)	Number of Enlisted Personnel
(classes)	(class frequencies)
\$ 50.00 - \$ 59.99	8
\$ 60.00 - \$ 69.99	10
\$ 70.00 - \$ 79.99	16
\$ 80.00 - \$ 89.99	14
\$ 90.00 - \$ 99.99	10
\$100.00 - \$109.99	5
\$110.00 - \$119.99	2
	Total 65

Table 2: Frequency Distribution of the Weekly Wages in Dollars of 65 Enlisted Personnel Aboard a Destroyer

With reference to Table 2 determine:

1. The lower limit of the sixth class. Ans. \$100.00

2. The upper limit of the fourth class. Ans. \$ 89.99

3. The class mark of the third class.
   Class mark of 3rd class = ½(\$70.00 + \$79.99) = \$74.995
   (For most practical purposes this is rounded to \$75.00)

4. The class boundaries of the fifth class.
   Lower class boundary of 5th class = ½(\$90.00 + \$89.99) = \$89.995
   Upper class boundary of 5th class = ½(\$99.99 + \$100.00) = \$99.995

5. The size of the fifth class interval
   Size of 5th class interval = upper boundary of 5th class - lower boundary of 5th class = \$99.995 - \$89.995 = \$10.00

HISTOGRAMS AND FREQUENCY POLYGONS

Histograms and frequency polygons are graphical representations of frequency distributions.

1. A histogram consists of a set of rectangles having:

   (a) Bases on a horizontal axis (the X axis) with centers at the class marks and lengths equal to the class sizes.

   (b) Areas proportional to class frequencies.

2. A frequency polygon is a line graph of class frequency plotted against class mark. It can be obtained by connecting midpoints of the tops of the rectangles in the histogram.

Figure 1 shows a histogram and frequency polygon developed from the data listed in Table 1.

RELATIVE FREQUENCY DISTRIBUTIONS

The relative frequency of a class is the frequency of the class divided by the total frequency of all classes, and is generally expressed as a percentage. For example, the relative frequency of the class 66 - 68 in Table 1 is 42/100 = 42%. The sum of the relative frequencies of all classes is 1, or 100%. If the frequencies in Figure 1 are replaced by corresponding relative frequencies, the figure is called a relative frequency distribution.

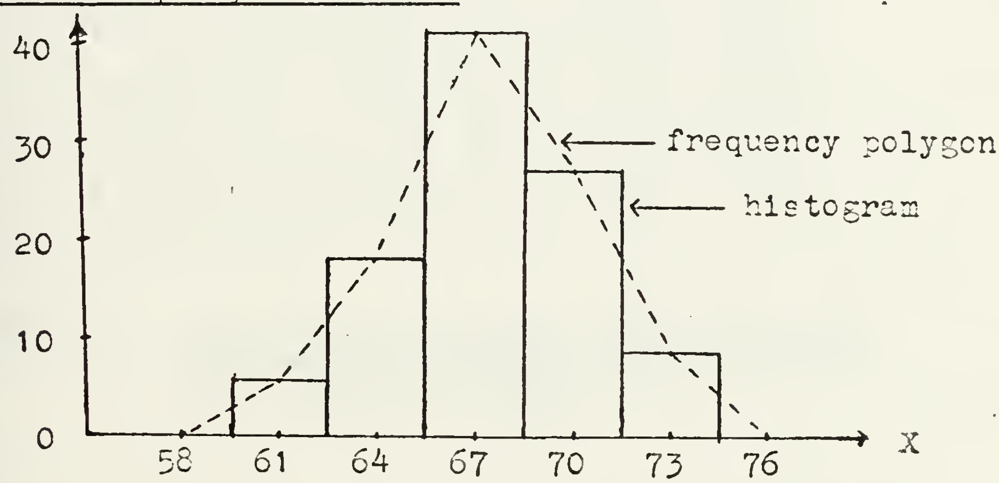


Figure 1: Histogram and Frequency Polygon of Data Listed in Table 1.

CUMULATIVE FREQUENCY DISTRIBUTIONS AND OGIVES

The total frequency of all values less than the upper class boundary of a given class interval is called the cumulative frequency up to and including that class interval. For example, the cumulative frequency up to and including the class interval 66 - 68 in Table 1 is 5 + 18 + 42 = 65, which means that 65 students have heights less than 68.5 inches. A table presenting such cumulative frequencies is called a cumulative frequency distribution, cumulative frequency table, or briefly a cumulative distribution, and is shown in Table 3 for the Naval Academy student height distribution. A graph showing the cumulative frequency less than any upper class boundary plotted against the upper class boundary is called a cumulative frequency polygon or ogive; an example is shown in Figure 2 for the Naval Academy student height distribution.

Height (inches)	Number of Students
less than 59.5	0
less than 62.5	5
less than 65.5	23
less than 68.5	65
less than 71.5	92
less than 74.5	100

Table 3: Cumulative Frequency Distribution of the Naval Academy Student Height Distribution of Table 1

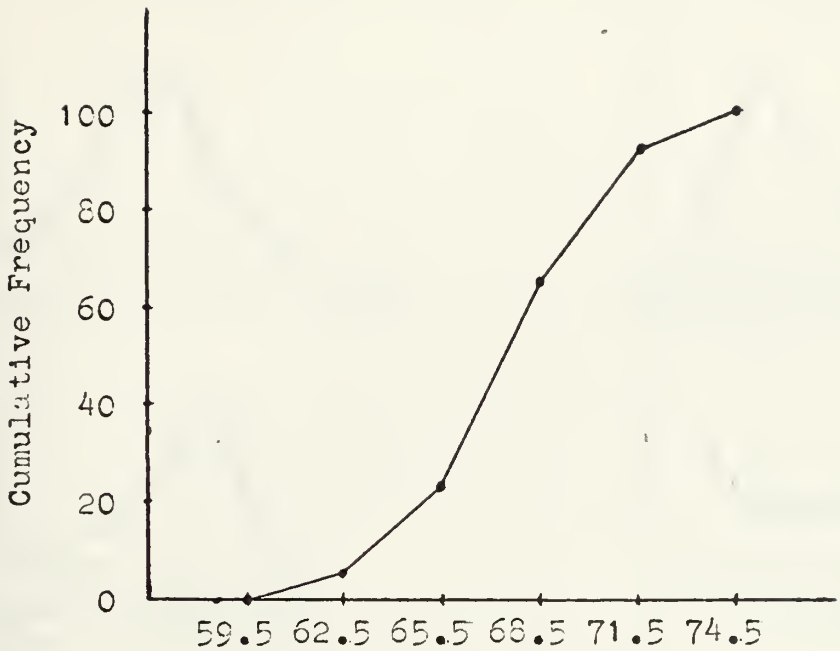


Figure 2: Ogive for the data of Table 3

RELATIVE CUMULATIVE FREQUENCY DISTRIBUTIONS

The <u>relative cumulative frequency</u> or <u>percentage cumulative frequency</u> is the cumulative frequency divided by the total frequency. For example, the relative cumulative frequency of heights of Naval Academy students with heights less than 68.5 inches is 65/100 = 65%, which indicates that 65% of the students have heights less than 68.5 inches.

TYPES OF FREQUENCY CURVES

Many times <u>relative</u> frequency curves are smoothed--that is, a smooth curve is "fitted" to a relative frequency polygon--and are used to represent situations arising in ASW applications, such as probability of detection. (These <u>probability distributions</u> are developed in Chapters IV through VI.)

Some of the more common frequency curves arising in ASW applications take on the characteristic shapes as indicated in Figure 3.

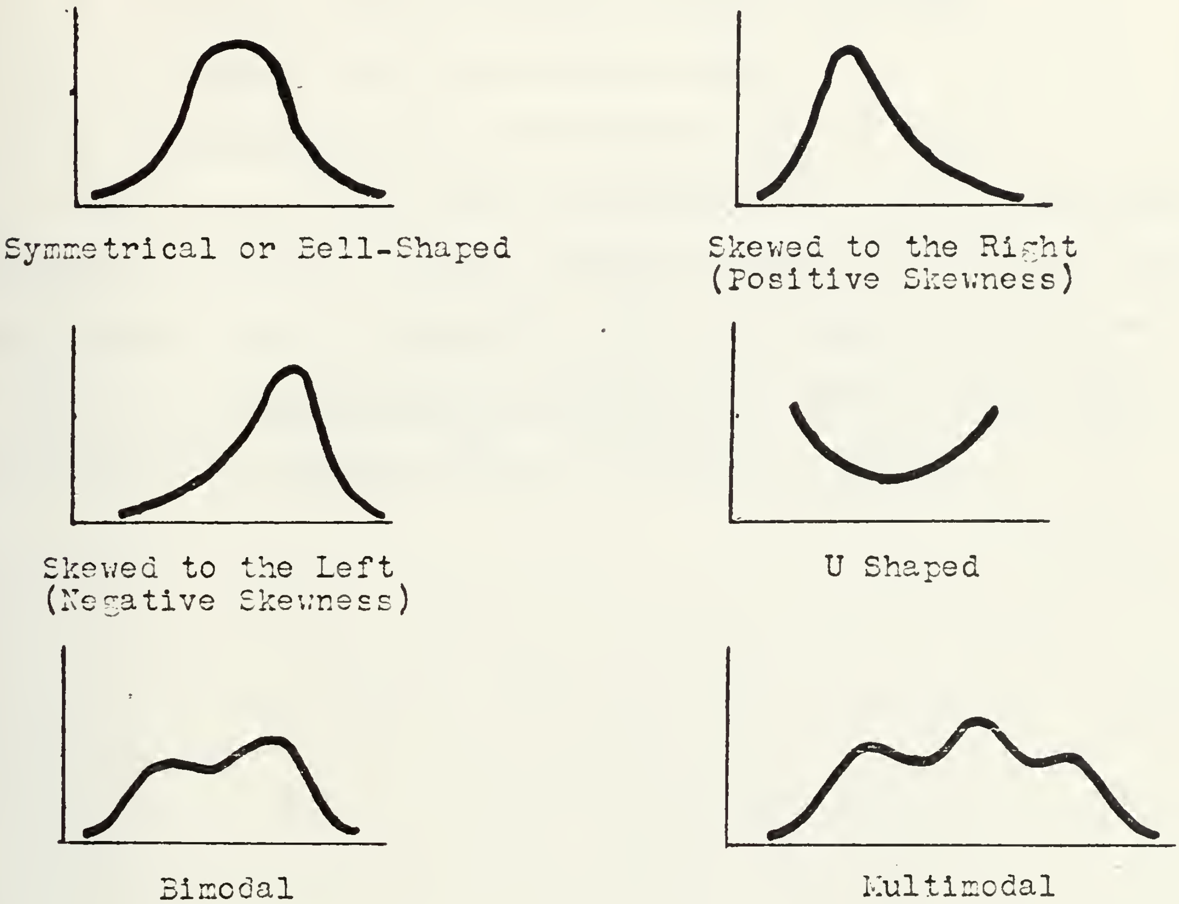


Figure 3: Types of Frequency Curves

1. The <u>symmetrical</u> frequency curves are characterized by the fact that observations equidistant from the central maximum have the same frequency. An important example is the Normal curve shown in Chapter VI.
2. In the <u>skewed</u> frequency curves the tail of the curve to one side of the central maximum is longer than that to the other. If the longer tail occurs to the right the curve is said to be skewed to the right or to have positive skewness, while if the reverse

is true the curve is said to be skewed to the left or to have negative skewness.

3. A U-shaped frequency curve has a maxima at both ends.

4. A bimodal frequency curve has two maxima.

5. A multimodal frequency curve has more than two maxima.

Supplementary solved problems covering arrays, frequency distributions, histograms, frequency polygons, cumulative frequency distributions, ogives, and frequency curves are presented at the end of Chapter 2 Schaum's Outline Series, Theory and Problems of Statistics. These problems should be consulted if a greater review of the material is desired.

CHAPTER II: SPECIAL MEASURES OF SAMPLE DATA: CENTAL TENDENCY AND DISPERSION

MEASURES

Although data relating to observations of events (for example, the number of submarine detections in a given region during a specific period of time) are important in developing methods for predicting similar events in the future, the data in the raw state are usually not directly suitalbe for making these predictions. The data must be reduced to a format that is still more manageable than the frequency histogram of Chapter I in order to facilitate examination further, and to a format from which the pattern of events on which the data are based can be determined. Such a representation may consist of "measures" of the data. Through the use of these measures, an entire data set can, effectively, be summarized by two or three single numbers. As a result, these measures are very important, and useful, in that they are devices by which sets of sample data can be quickly described.

Two particularly useful measures of sample data are those which indicate or describe a relative central value of the sample data (central tendency), and the degree to which the data are spread out about a central value (dispersion). A measure of central tendency can be utilized to indicate a "common quantity" which can be used to describe roughly the entire sample (and possibly even the entire population). And a measure of dispersion can be used as an indicator of the relative "error" or "spread" that exists when a measure of central tendency is used to describe the sample or the population.

A digression will be made at this point to introduce two useful forms of notation, index or subscript notation, and summation notation.

SUBSCRIPT NOTATION

When it is necessary to describe a variable, X, and the values it can assume, and also to distinguish between individual members of the group (these members are specific values which the variable X can assume—the 1st x, the 2nd x, etc.), subscript notation is a concise, useful means of describing specific values of the variable. To describe the values of the variable X with subscript notation, the first value (1st x) of the group is denoted by x_1, the second value by x_2 and so forth to the last, say the n^{th}, value by x_n. The entire group of values of the variable, X, can then by denoted by

$$X = (x_1, x_2, \dots , x_n),$$

where the three dots indicate that middle values are included in the set, but are not explicitly written down for brevity.

SUMMATION NOTATION

Summation notation is a concise and useful means of denoting the sum of all or part of the individual observed values of a variable X. The notation used to denote the sum

$$x_1 + x_2 + x_3 + \dots + x_n$$

is

$$\sum_{i=1}^{n} x_i \ .$$

The capital Greek letter Σ (sigma) denotes that a sum is to be taken over

the observed values of the variable x_i which follow it, and the values of x_i that are to be summed are those with subscripts between $i = 1$ and $i = n$. For example, to sum the third through the fifth observed values of X, the summation is written as

$$\sum_{i=3}^{5} x_i \ .$$

That is,

$$\sum_{i=3}^{5} x_i = x_3 + x_4 + x_5 \ .$$

This example shows that the lower limit on i need not be 1. Further, if the values X can assume (say, a total of six values) are $x_1 = 0$, $x_2 = 3$, $x_3 = 7$, $x_4 = 2$, $x_5 = 1$, and $x_6 = 9$, then

$$\sum_{i=3}^{5} x_i = 7 + 2 + 1 = 10 \ .$$

The following example shows how this notation can be used in an ASW setting. Suppose there are a total of six regions each of which can be monitored by a single sensor, where the number of possible targets in each region is known for given periods of time (from intelligence). It is important for the search planner to know the number of potential detection opportunities available from the monitored regions so that he can ensure that sufficient localizing/tracking platforms (ships, aircraft, submarines) are on standby for employment upon initial detection. As the decision to monitor regions three through five has been previously made, he wishes to specify the number of detection opportunities available in these regions on a given day in a standard format regardless of the particular values entered in the format for a specific period; this is because he wishes to include this calculation in a larger computer

program. The notation

$$\sum_{i=3}^{5} x_i ,$$

is a convenient format for this purpose. In any particular run of the program, numerical values for x_3, x_4 and x_5 must be entered.

An important property of summation is that if a constant value, k, is summed over a range of N indices, then the summation of this quantity is equivalent to multiplying the constant k by N, since, for example,

$$\sum_{i=2}^{5} K = \underbrace{K + K + K + K}_{} = 4K$$

$$N = (5 - 2) + 1 = 4 \text{ terms}$$

MEAN, MEDIAN, AND MODE

Three particularly useful <u>measures</u> of <u>central tendency</u> are the <u>mean</u>, the <u>median</u>, and the <u>mode</u>.

The mean, which is similar to the commonly used notion of arithmetic average, is computed by adding the observed values of sample data and dividing this sum by the total number (N) of data points in the sample. The mean can be expressed in summation notation as

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} x_i ,$$

where N is the total number of observed values summed, and X is read as "the mean value of X." (Note that the upper limit on the summation index equals the reciprocal of the constant in front of the summation.) For example, the mean of the observed values $x_i = 2$, $x_2 = 5$, and $x_3 = 1$ of x is

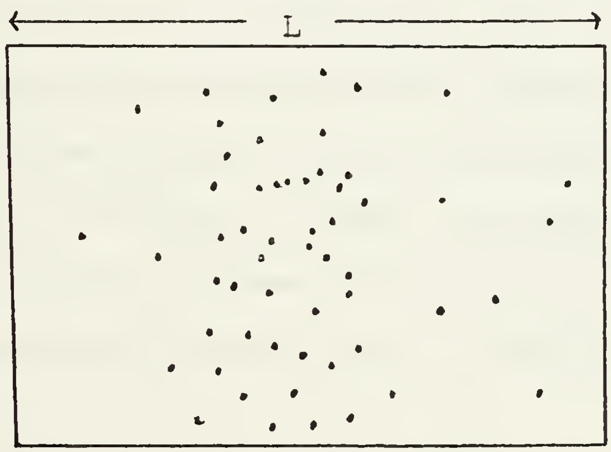
$$\bar{X} = 1/3 \; (\sum_{i=1}^{3} x_i) = 1/3 \; (2 + 5 + 1) = 8/3 .$$

The utility of this value ($\bar{X}$) for X is that it can be taken, in many cases, to state a "middle value" of the observed data, and is often used as the best single number estimate of the values of X.

For example, suppose it is necessary to search for an enemy submarine which is currently maintaining a "known" missile launching station, and the following conditions exist:

1. the "known" missile launch station is actually a large ocean region;
2. the area that can be searched effectively is one-half the size of the missile station;
3. the positions of previous enemy submarine detections within the missile launch area are available to the search planner.

The above information is depicted in Figure 4a below, where the dots represent the previous detection positions.



Missile Launch Station

Figure 4a

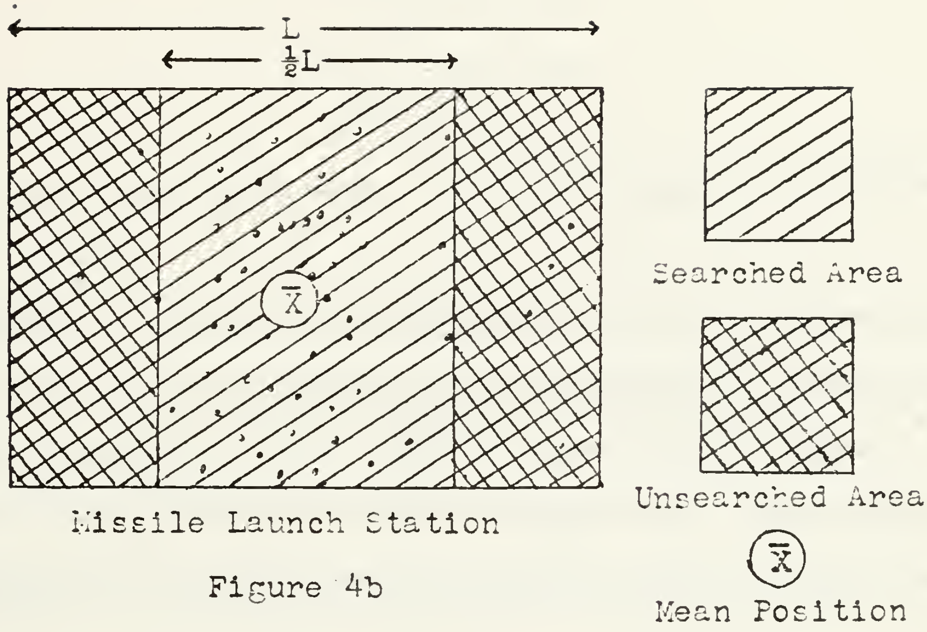


Figure 4b

Figure 4: Submarine Search Information:
a. Positions of Previous Detections;
b. Mean Search Area

It can be seen from Figure 4a that the positions tend to cluster about a certain small enclosed portion of the launch region; therefore, it seems reasonable to search about this small area in order to increase efficiency. The center of this small region is the mean position of previous enemy submarine contacts. Therefore, to obtain the position about which the search should be centered, the mean of the previous positions is calculated. The search is subsequently centered about this mean position, as seen in Figure 4b.

A special form of the mean is the weighted (arithmetic) mean. This form allows one to account for differing "levels of importance" of the individual sample observations. For example, if in a three element sample, x_1 were considered to be twice as important as x_2, and x_3 four times as important as x_2, the weighted mean of the sample would be computed as

$$\bar{X} = \frac{m_1x_1 + m_2x_2 + m_3x_3}{m_1 + m_2 + m_3}$$

$$= \frac{2x_1 + 1x_2 + 4x_3}{2 + 1 = 4}$$

where the m_i is the relative weight, or level of importance, assigned to the i^{th} element.

The <u>median</u> value is the value which exactly divides the sample into two equal parts; there are exactly as many observations greater than the <u>median</u> value as there are less. When the number of observations is even, the median is exactly half-way between the <u>two</u> "center" observations, but for an odd number of observations the median value <u>is</u> the center observation.

For example, given the sample

$$x_1 = 2,\ x_2 = 5,\ x_3 = 1,\ x_4 = 7$$

when arrayed is 1, 2, 5, 7 and has as its median

$$\tilde{X} = \frac{2 + 5}{2} = 3.5\ ,$$

and we note that this is not one of the values actually observed.

The <u>mode</u> or <u>modal value</u> of the sample is the most frequently occurring value/s in the sample. If more than one value occurs at this same maximum frequency (of occurrence), then the sample is said to be <u>multi-modal</u>; such a sample would have a frequency curve with more than one peak. For example, for the sample $x_1 = 2$, $x_2 = 5$, $x_3 = 2$, and $x_4 = 3$, the mode is $\hat{X} = 2$, since it occurs twice in the sample, while all other observed values appear once. Although the mode is easy to ascertain from a sample, it is probably the least useful of the measures of central tendency treated here.

It is seen that each of the three measures of central tendency mentioned can be used as an estimate of (or statistic which describes) the sampled population; however, in fact, the mean is probably the most useful of the three estimators--this is so because the mean is less biased than the other two measures. In other words, the mean is not influenced by extreme sample values to the extent the median and mode are. Therefore, for most practical applications, the mean of a sample reflects the central tendency of the parent population more accurately than the median or the mode.

RANGE, STANDARD DEVIATION AND VARIANCE--MEASURES OF DISPERSION

A useful statistical indicator on sample data and its resident population is the spread of the sample points (usually expressed relative to, or about, the mean). Such an indicator is called a measure of dispersion (spread).

Three particularly useful measures of dispersion are the range, the standard deviation, and the variance.

The range is obtained by taking the difference between the largest and the smallest observed value in the sample. While the range of sampled data is easy to compute, it is not as accurate an indicator as the other two mentioned. This is due to the large degree to which one or two sample values (a low frequency of occurrence) at relatively extreme distances from the center of the greatest portion of the sample can bias the estimate of the actual spread (or dispersion) of the sampled population.

The standard deviation σ (lower case Greek sigma) is obtained by first computing the sample mean, and then using the formula

$$\sigma = \sqrt{\left[\frac{1}{N} \sum_{i=1}^{n} (x_i - \bar{X})^2\right]} .$$

Since the standard deviation tends to <u>weight</u> each observed value by its frequency of occurrence, it is seen to be a more accurate measure of dispersion than the range; that is, the standard deviation is not unduly biased by a few extreme sample values.

The <u>variance</u> is found by simply squaring the standard deviation. Consequently, the variance has all the advantages of the standard deviation. The mathematical form for the <u>variance</u> is

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{n} (x_i - \bar{X})^2 .$$

An equivalent form for the variance is

$$\sigma^2 = \overline{X^2} - \bar{X}^2 ,$$

where

$$\overline{X^2} = \frac{1}{N} \sum_{i=1}^{n} x_i^2 \qquad \text{and } \bar{X}^2 = \left(\frac{1}{N} \sum_{i=1}^{n} x_i\right)^2 .$$

The latter equation is the "short method" for computing σ^2 and is the recommended method of computing the variance; the positive square root of this equation is the "short method" for computing σ and is recommended for computing the standard deviation.

As an example which illustrates the equivalence of the two listed forms for computing σ^2, consider the case where three BT (bathythermograph) readings are observed and the first shows a sea surface temperature of 50°F, the second 48°F, and the third 52°F. The variance for these readings is first found by

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{n} (x_i - \bar{X})^2 = 1/3 \sum_{i=1}^{3} (x_i - \bar{X})^2$$

$$= 1/3 \left[(50 - \bar{X})^2 + (48 - \bar{X})^2 + (52 - \bar{X})^2 \right]$$

$$\sigma^2 = 1/3 \ (0 + 4 + 4) = 8/3 \ .$$

Now, σ^2 is found by

$$\sigma^2 = \overline{X^2} - \bar{X}^2 \ ,$$

where

$$\overline{X^2} = 1/3 \sum_{i=1}^{3} x_i^2 = 1/3 \ (50^2 + 48^2 + 52^2) = 7508/3$$

and

$$\bar{X}^2 = (1/3 \sum_{i=1}^{3} x_i)^2 = \left[1/3 \ (50 + 48 + 52) \right]^2$$

$$= 50^2 = 7500/3,$$

so that

$$\sigma^2 = 7508/3 - 7500/3 = 8/3 \ .$$

(The second method looks longer than the first, but the mean had to be calculated prior to the first method.) Hence, this example illustrates the equivalence of the two formulas for σ^2.

As an example of the usefulness of the standard deviation, consider the Normal probability frequency distribution (which is covered in detail in Chapter VI) shown in Figure 5. If sigma (the standard deviation) is known for the distribution, the percentage of cases in which any (observed) value could be expected to have come from the intervals (of the distribution) $\bar{X} - \sigma$ to $\bar{X} + \sigma$, $\bar{X} - 2\sigma$ to $\bar{X} + 2\sigma$ or $\bar{X} - 3\sigma$ to $\bar{X} + 3\sigma$ is 68.27%, 95.45%, or 99.73%, respectively, as shown.

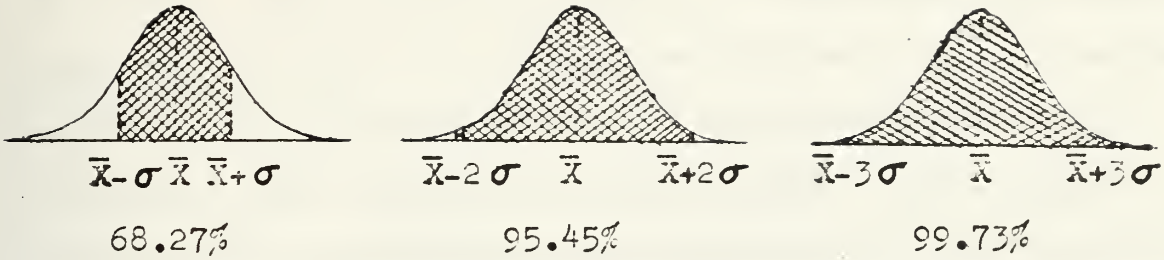


Figure 5: Normal Probability Frequency Distributions

As an application, consider a situation where it is necessary to obtain the probability of detecting a submarine at a distance of 45 n.m. from the detection device (which is denoted by P_{det}(45 n.m.)), given the FOM (Figure-of-Merit), the pertinent PLP (Propagation Loss Profile), and the σ-value for the probability of detection.

The first step is to visualize a Normal proabbility curve whose "x-axis" is located vertically along the 45 n.m. line of the PLP, and whose mean value is on the FOM line along the "x-axis," as seen in Figure 6.

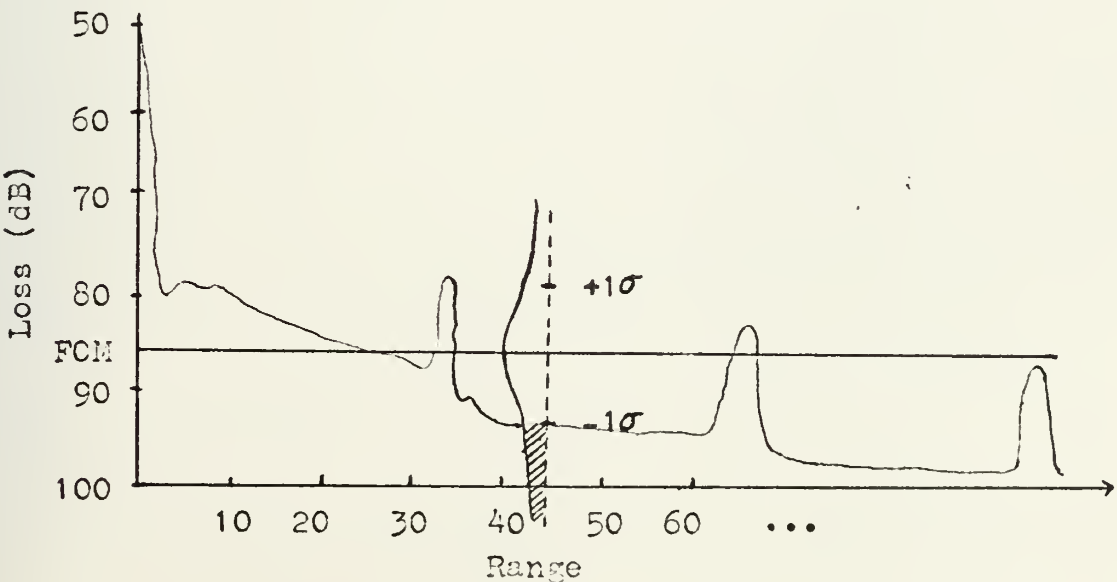


Figure 6: PLP (Propagation Loss Profile)

Next, recall that the area under the Normal curve to either side of the FOM (the mean value for the distribution) corresponds to 50% probability of detection. Note that the P_{det}(45 n.m.) is represented by the darkened area under the Normal curve and under the PLP curve at 45 n.m., and that the intersection of the PLP and the axis of the probability curve is at -95 dB or -1σ relative to the mean (the FOM). Now the P_{det}(45 n.m.) is calculated as the difference between 50% (the lower-half of the area under the Normal curve) and one-half of the area (under the Normal curve) from -1σ to $+1\sigma$ (which was previously shown to be 68.27%), or

$$P_{det}(45 \text{ n.m.}) = \left[50 - \tfrac{1}{2}(68.27)\right]\% = 15.87\% .$$

(Other methods for obtaining probabilities are presented in Chapters IV through VI.)

Supplementary solved problems covering special measures of sample data are presented at the end of Chapters 3 and 4 of Schaum's Outline Series, Theory and Problems of Statistics. These problems should be consulted for a further review of the material.

CHAPTER III: COMBINATORIAL ANALYSIS

COMBINATORIAL ANALYSIS

Combinatorial analysis involves application of counting principles to obtain probabilities of complex events. These basic principles are presented in this chapter.

THE FUNDAMENTAL COUNTING PRINCIPLE

If an event can happen in any of n_1 ways, and if when this has occurred another event can happen in any of n_2 ways, then the number of ways in which both events can happen is n_1n_2. Example: if 3 submarines are to transit 2 barriers, then the number of ways the submarines can transit the barriers are (2)(3) = 6 ways; if 5 submarines are to transit 3 barriers then there would be (3)(5) = 15 ways.

FACTORIAL n

Factorial n, denoted by n!, is equal to n(n - 1)(n - 2)...1. Thus 5! = (5)(4)(3)(2)(1) = 120, 4!3! = ((4)(3)(2)(1))((3)(2)(1)) = 144. It is convenient to define 0! = 1. Further, each factor in the factorial product must be a (nonnegative) integer.

PERMUTATION PRINCIPLE

A permutation is an arrangement of r out of n objects with attention given to the order of arrangement. The number of permutations of n objects

taken r at a time is denoted by nPr, P(n,r) or Pn,r and is given by

$$nPr = n(n - 1)(n - 2)\ldots(n - r + 1) = \frac{n!}{(n - r)!}$$

For example, how many ways may 5 destroyers be placed in a 5-position screen? The first position can be occupied by any one of the 5 destroyers; that is, there are 5 ways of filling the first position. When this has been done there are 4 ways of filling the second position. Then there are 3 ways of filling the third position, 2 ways of filling the fourth position and 1 way of filling the last position. Therefore, n represents the 5 destroyers and r represents the number of destroyers that may be used in the screen which also is 5; hence,

$$nPr = \frac{n!}{(n - r)!} = \frac{5!}{(5 - 5)!} = \frac{5!}{0!} = \frac{5!}{1} = 5! = (5)(4)(3)(2)(1) = 120$$

Now suppose 5 destroyers are available for placement in a 3 position screen. How many ways may the destroyers be placed in the screen? The first position can be occupied by any one of the 5 destroyers, then the second position can be occupied by any of the remaining 4 destroyers and finally the last position may be occupied by any of the remaining 3 destroyers. Therefore, n represents the number of destroyers which is 5, and r represents the number of destroyers that may be used which is 3; so that,

$$nPr = 5P3 = \frac{n!}{(n - r)!} = \frac{5!}{(5 - 3)!} = \frac{5!}{2!} = \frac{(5)(4)(3)(2)(1)}{(2)(1)} = \frac{120}{2} = 60.$$

COMBINATION PRINCIPLE

A combination is a selection of r out of n objects <u>without</u> counting the order of arrangement. The number of combinations of n objects taken

r at a time is denoted by nCr, C(n,r), Cn,r or $\binom{n}{r}$ and is given by

$$nCr = \frac{n(n-1)\ \dots\ (n-r+1)}{r!} = \frac{n!}{r!(n-r)!} = \frac{nPr}{r!} \quad .$$

For example: Suppose a submarine with six torpedoes approaches a convoy of 10 ships. If the ships are selected at random and one torpedo is shot at each ship, how many ways may the torpedoes be shot at the 10 ships?

$$nCr = \frac{n!}{r!(n-r)!} = \frac{10!}{6!(10-6)!} = \frac{(10)(9)(8)(7)(6)(5)(4)(3)(2)(1)}{((6)(5)(4)(3)(2)(1))((4)(3)(2)(1))} =$$

$$\frac{3,628,800}{(720)(24)} = \frac{3,628,800}{17,280} = 210 \text{ ways}$$

Now suppose 3 of the ships are <u>warships</u> protecting the convoy and it is desired to keep clear of <u>these</u> ships. If the remaining 7 ships are selected at random and one torpedo is shot at each ship, how many ways may the torpedoes be shot at the 7 ships?

$$nCr = \frac{n!}{r!(n-r)!} = \frac{7!}{6!(7-6)!} = 7 \text{ ways}$$

Supplementary solved problems covering permutations and combinations are presented at the end of Chapter VI of Schaum's Outline Series, <u>Theory and Problems of Statistics.</u> These problems should be consulted for further review of the material.

CHAPTER IV: ELEMENTARY PROBABILITY THEORY

ELEMENTARY PROBABILITY THEORY

The concept of probability is commonly used as a means with which to state (or predict) the percentage of possible cases in which a specific event will occur; for example, the percentage of events in which a specific number will appear on a single toss of a die. The probability of such an event occurring is, in many cases, given as the ratio of the number of ways in which the event can occur to the number of ways in which ALL events can occur. So the probability that a two, for example, will occur on a single toss of a fair die is

$$P(2) = \frac{1}{6} ,$$

where the numerator represents the one way in which a two can occur, and the denominator represents the six ways in which each of the numbers 1, 2, 3, 4, 5, 6 can occur; hence, the probability that a two will be the outcome of a single toss of a die is one-sixth, or in other words, the "chances" of this outcome occurring--an event--are "one in six."

Another method of specifying probability of an event is to state the probability in terms of proportional areas. The basis of this method is to think of the probability of occurrence of ALL events as an area, say "A," and the probability of a particular event is simply the ratio of the areas

$$P(\text{event of interest}) = \frac{B}{A} ,$$

where B is the area corresponding to the event of interest.

Probabilities specified as proportional areas can be represented by diagrams of the corresponding regions. A diagram of two regions with areas "A" and "B" is shown in Figure 7.

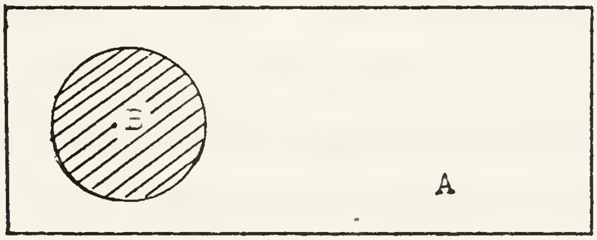


Figure 7: Proportional Areas for Probability Analogy

This figure facilitates understanding the probability of event "B" by illustrating the area which is <u>assigned</u> to a region relative to the total area, "A," assigned to the occurrence of ALL events. Since the diagram visually represents the probability of the event, it is a useful means of <u>intuitively</u> developing the concept of probability. (A carefully cultured intuition can be very useful when dealing with probability.)

It is necessary here to define some terms which will make the formulation of basic probability theory easier. The first of these terms is PROBABILITY EXPERIMENT, which is defined as a type of trial procedure for selecting a specific outcome. An example is that of tossing a die to see what number comes up and noting the occurrences (or of flipping a coin to see whether "heads" or "tails" occurs).

Next, a SAMPLE SPACE is the totality of outcomes which can occur for the probability experiment. In terms of plane regions, the sample space is the large region ("A" in the previous example) which contains all possible outcomes of a specific probability experiment.

An OUTCOME is a specific value (or set of values) resulting from a single iteration of a probability experiment, and it is the basic element (or smallest unit) of the sample space (for the specific probability experiment). An example would be the occurrence of one specific permutation (of the 216 possible permutations) of the numbers appearing on the three die in the (probability) experiment of tossing three dice on a single toss and noting the numbers which appear, where the goal of the experiment is to see whether or not a one, a two, and a three appear on the toss. This can be represented as one of the totality of points that make up the sample space (a three dimensional space, as each outcome is composed of three numbers).

An EVENT, on the other hand, is a set of outcomes; in the previous example, an outcome is the sequence of the three specific numbers which occur as a result of the toss of the three dice, and the event consists of six points of interest (outcomes) in 3-space, as shown in Table 4.

Die #1	Die #2	Die #3
(1 ,	2 ,	3)
(1 ,	3 ,	2)
(2 ,	1 ,	3)
(2 ,	3 ,	1)
(3 ,	1 ,	2)
(3 ,	2 ,	1)

Table 4: The Six Outcomes That Comprise (Satisfy) the Event of Obtaining a One, a Two, and a Three on a Single Toss of Three Dice

In terms of the proportional areas analogy, an event is an area ("B" in the earlier example) which contains one or more outcomes and is a subarea of the sample space ("A). An extension of this terminology is the term JOINT EVENTS, which describes a combined event in which two or more simple events occur on the same iteration of the experiment.

In summary, the sample space is comprised of all possible outcomes of a probability experiment, and these outcomes are grouped into sets called events--outcomes are the elements of an event.

The remaining basic concept in the theory is that of understanding the PROBABILITY FUNCTION. A probability function is a device by which a number (or area) in the sample space is assigned (as previously mentioned) to an event. An example of a probability function is the assignment of the value 1/3 to the probability of each of the events that a single submarine, during a single transit, will cross one of three barriers each of which the submarine is equally-likely to cross. This is written as

$$P(\text{SS transits \#1}) = 1/3$$

$$P(\text{SS transits \#2}) = 1/3$$

$$P(\text{SS transits \#3}) = 1/3 \ ,$$

which can also be represented in an area diagram as in Figure 8, where it is seen that all three areas are equal and that no area overlaps any other, which is interpreted to mean that the submarine can cross only one of the barriers during a single transit.

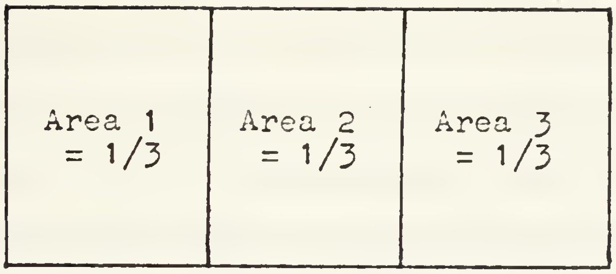


Figure 8: Proportional Area Depiction of Equal-likelihood Probability Barriers

There are basic restrictions on the probability function. One of these is that the values assigned to events by the probability function must sum to one, or

$$\sum_{i=1}^{n} p_i = 1 \ ,$$

where i corresponds to the i^{th} event.

Another is that the values assigned must be between zero and one, or

$$0 \leq p_i \leq 1 \ .$$

This restriction is a result of the fact that probability is a ratio of a small area to a larger one. Also, since a specific event in a given experiment will either <u>occur</u> or <u>not occur</u>, the sum of the probabilities that the event will and the event will not occur must equal one. This is represented mathematically as

$$P(\text{Specific event occurs}) + P(\text{Specific event does not occur}) = 1$$

or by rearrangement

$$P(\text{Specific event does not occur}) = 1 - P(\text{Specific event occurs})$$

or

$$P(\text{Specific event occurs}) = 1 - P(\text{Specific event does not occur}) \ .$$

As seen by the last two arrangements of the first equation, the probability of an event occurring and the probability of the same event not occurring on the same trial of an experiment are complements of each other in the probability sense; that is, the probability of an event occurring can also be described by one minus the probability of the same event not occurring. As will be seen later, this idea of PROBABILITY COMPLEMENT of the occurrence of an event has many useful computational applications.

We now pass from probabilities for single events to probabilities for several events. There are two cases of interest: the probability of joint events, and the probability of at least one of two (or more) events. Joint events are denoted by the word "and"; for example, the probability of events A and B occuring. The other case is denoted by the word "or"; for example, the probability of event A or B (or both occurring, where A or B or both is the sense in which it is used in this course). This "or" probability is shown by the proportional areas analogy as in Figure 9.

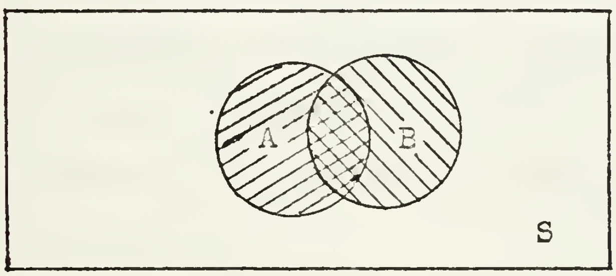


Figure 9: Probability of Events A "or" B

It is seen from this figure that the probability of event A or B is represented as the probability of A plus the probability of B, minus the probability of A and B occurring. The reason for subtracting the joint probability of A and B is that the area corresponding to the joint probability ("cross-hatched" in the figure) is common to both event A and

event B, so that when area A is added to area B the "cross-hatched" area is added <u>twice</u>; therefore, the joint probability must subsequently be subtracted once. The probability of A <u>or</u> B occurring is therefore expressed mathematically as

$$\begin{aligned} P(A \text{ or } B) = P(A + B) &= P(A) + P(B) - P(A \text{ and } B) \\ &= P(A) + P(B) - P(A \cdot B). \end{aligned}$$

Figure 10 shows a case in which events A and B have no area of overlap; that is, A and B do not have any area in common.

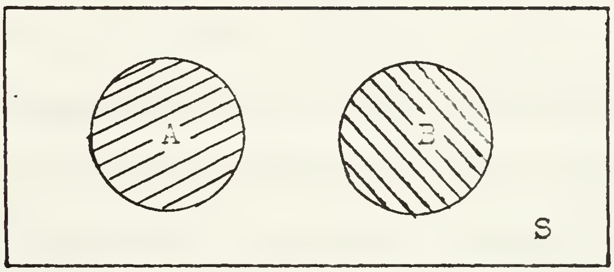


Figure 10: Probability of Events A "or" B when Events A and B are Mutually Exclusive

It is easily seen from this figure that any outcome in event A cannot possibly also be a part of event B; in fact, the existence of the outcome in <u>either</u> A or B <u>precludes</u> its simultaneous existence in the other event. Two such events are said to be MUTUALLY EXCLUSIVE.

The joint probability of A and B can be shown through the proportional areas analogy as in Figure 11.

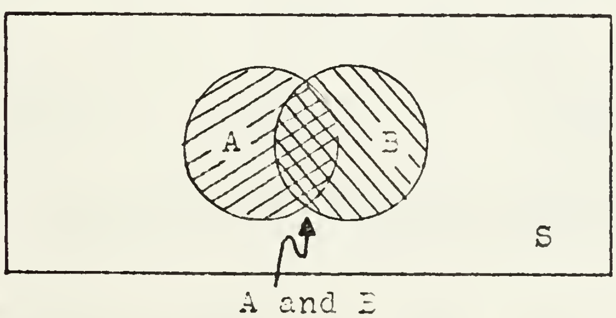


Figure 11: The Joint Probability of Events A and B

It is seen from the figure that the probability of the joint event is the ratio of the "cross-hatched" area (the joint event itself) to the entire area (which is equal to one since it is a probability space). Mathematically the joint probability of events A and B is represented as

$$P(A \text{ and } B) = P(A \cdot B) = P(A) + P(B) - P(A + B) \ .$$

(This is simply a rearrangement of the "or" probability equation developed previously.)

The symbol P(B A) denotes the CONDITIONAL PROBABILITY of event B <u>given</u> that event A has already occurred; that is, it is the probability of B <u>on the condition</u> that event A has occurred previously. It is seen from Figure 12, that P(B A) differs from P(B) by the fact that the sample space (S) from which B is drawn is larger than the resultant sample space (S') from which B is drawn when A has been given to occur on this same outcome (the event "B given A"). This is due to the fact that S has been <u>reduced</u> to S' <u>after</u> it is known that the event yet to occur (B) is contained in the event that has already occurred (A); therefore, in terms of the ratio idea, the denominator has been decreased, thereby <u>increasing</u> the probability of B for the conditional case (relative to the unconditioned case).

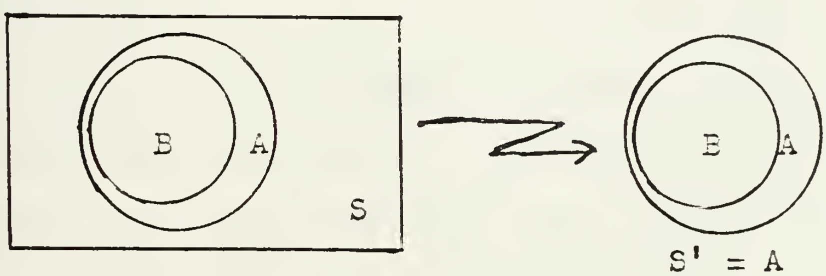


Figure 12: Comparison of Unconditioned Probability to Conditional (or Conditioned) Probability

CONDITIONAL PROBABILITY actually epitomizes the usefulness of probability to ASW. That is, generally, an area is searched for the presence of a

submarine, and both the cases in which the submarine 1) has or 2) has not been detected as a result of previous search efforts are important information. Each allows the searcher to refine his prediction of the target's position (in most cases) to a smaller and smaller portion of the original search region and, thereby, increase the probability of detecting the submarine on subsequent portions of the search. It is obvious how this occurs in the case where the target is detected on a given previous effort, but the case of no detection on previous search efforts may not be so obvious. The fact that the submarine was not detected on previous portions of the search allows the searcher to eliminate (again, in most cases) that portion which has been searched--similar in idea to "sterilization"--thereby conditioning the original sample space (this should be visualized as multiple iterations of the process depicted in Figure 12). Having previous knowledge (intelligence data) which allows the searcher to eliminate portions of the (original) search area is another way in which this conditioning can occur.

The joint probability of events A and B (discussed previously) can be represented as

$$P(A \text{ and } B) = P(B|A) \cdot P(A) \ .$$

This equation has more intuitive appeal than the previous equation for joint probability, as it is simply an extension of the "Counting Principle"; it indicates that the probability of both event A and B occurring is the probability that event A occurs and "event" B given A occurs.

From the last equation it is seen that if P(A)   0, then the conditional probability of B given A is mathematically represented as

$$P(B|A) = P(A \cdot B) \ / \ P(A) \ .$$

Three important methods for computing conditional probabilities are:

1. By utilizing the definition of conditional probability

$$P(B|A) = P(A \cdot B) \ / \ P(A) \ .$$

For example, it is known (by some means or another) that a submarine skipper will <u>not</u> choose one of three known and equally-likely barriers to transit, then the resulting probability is <u>increased</u> beyond that for the original choice due to the sample space drawn on having been reduced from three possible paths to two. If the barrier eliminated is #1, then the resulting sample space S' consists of barriers #2 and #3, vice the original sample space S which was made up of barriers #1, #2, and #3. Therefore, the probability that the submarine will transit barrier #2, given it will transit either #2 or #3 is seen to be

$$P(B|A) = P(A \cdot B) = \frac{P\left[(\#2 \text{ or } \#3) \text{ and } (\#2)\right]}{P(\#2 \text{ or } \#3)}$$

$$= \frac{P(\#2)}{P(\#2 \text{ or } \#3)} \quad \text{since the only way in which \#2 \underline{and} (\#2 \underline{or} \#3) can occur is if \#2 occurs}$$

$$= \frac{1/3}{1/3 + 1/3} = \frac{1/3}{2/3} = 1/2 \ .$$

2. By simply writing the conditional probability down, as sometimes the information that A has occurred enables the determination of the probability of B.

For example, in the experiment of drawing two cards from a deck of playing cards one at a time without replacing the drawn card in the deck, it is known that the first card drawn is a Spade. Given that the first card is a Spade, the probability that the second card is also a Spade can immediately be written down as

$$P(\text{draw \#2 is a Spade draw \#1 is a Spade}) = 12/51 \ ,$$

since the original 13 Spades in the deck have been reduced to 12, and the original 52 total cards (the sample space) to 51 by the first draw.

The probability thus formed is consistent with the ratio concept of probability.

3. By utilizing Bayes' law, which will now be presented without proof and with only the following restriction.

Bayes' law holds for experiments on sample spaces which consist of definite groups of events that do not "overlap." Such a sample space is seen to correspond to the delimited area/s in Figure 13.

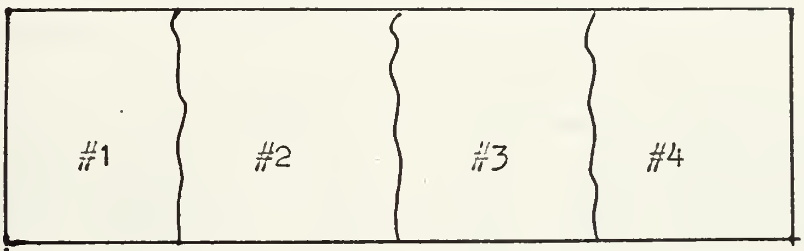


Figure 13: A Partitioned Sample Space

Such a sample space is said to be PARTITIONED.

Bayes' law states that on a partitioned space the probability of an event A occurring, given that event B has previously occurred is,

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)} ,$$

where $P(A^c)$ is the probability complement (covered earlier) of P(A), or $P(A^c) = 1 - P(A)$.

For example, suppose it is known from intelligence reports that missiles of type C and D are carried by exactly two classes of Soviet submarines, say class A and class B. Submarines of class A are known to carry 16 missiles of type C and eight of type D, while those of class B are known to carry ten of each. A missile is fired from a submarine and observed to be of type C. Find the probability that the missile was fired from a class A sub, under the following conditions:

1. Each missile to be fired is chosen at random from the total supply of missiles aboard the submarine, and

2. the Soviets normally operate twice as many subs of class B as of class A in the area.

The problem is to find P(A C); it can be found using Bayes' law. To determine the values to be used, it is best to first draw two blocks representing the class A and B submarines and the types of missiles they carry as shown in Figure 14. (Remember that there are two class B subs for each class A.)

Class A submarine	Two Class B submarines
16 C missiles 8 D missiles 24 missiles	10 C missiles 10 D missiles 20 missiles

Figure 14

By Bayes' law then

$$P(A|C) = \frac{P(C|A) \cdot P(A)}{P(C|A) \cdot P(A) + P(C|A^c) \cdot P(A^c)} ,$$

here $A^c = B$; therefore,

$$= \frac{P(C|A) \cdot P(A)}{P(C|A) \cdot P(A) + P(C|B) \cdot P(B)}$$

$$= \frac{(16/24) \cdot (1/3)}{(16/24) \cdot (1/3) + (10/20) \cdot (2/3)} = \frac{2/9}{5/9} = 2/5 .$$

As a second example of the application of Bayes' law, consider a circuit breaker used to protect a certain piece of ASW hardware from short circuits. At the moment the "on" switch is thrown to activate the equipment, the circuit breaker essentially "tests" the circuit for a short. If a short is detected, the circuit breaker should immediately trip; otherwise it should allow operation to continue.

Suppose it is known that although short circuits do not <u>always</u> trip the circuit breaker, they do trip it 99.5% of the time. On the other hand, suppose the probability that the circuit breaker trips when there is no short is 3%. If the probability that there is a short in the equipment when the switch is thrown is 0.01, find the conditional probability that there is no short circuit given that the circuit breaker does not trip.

The technique of solution involves several steps: 1) describe mathematically the probability "events" involved; 2) formulate what is given and immediate consequences of the given; 3) formulate the unknown; 4) write down Bayes' law as it applies to the events of this problem; 5) substitute the given numbers and solve.

1) Let: T be the event that the circuit breaker trips;

T^c the event that the circuit breaker does not trip

S the event that there is a short circuit when the "on" switch is thrown;

S^c the event that there is no short when the "on" switch is thrown.

2) We are given that

$P(T|S) = 0.995$, $P(T|S^c) = 0.03$, $P(S) = 0.01$.

From the given it follows immediately that

$P(T^c|S) = 0.005$, $P(T^c|S^c) = 0.97$, $P(S^c) = 0.99$.

3) The problem is to find

$$P(S^c|T^c) .$$

4) By Bayes' law,

$$P(S^c|T^c) = \frac{P(T^c|S^c) \cdot P(S^c)}{P(T^c|S^c) \cdot P(S^c) + P(T^c|S) \cdot P(S)} .$$

5) Substitute:

$$P(S^c | T^c) = \frac{(0.97)\ (0.99)}{(0.97)\ (0.99) + (0.005)\ (0.01)}$$

$$= \frac{0.9603}{0.9603 + 0.00005} = .999948$$

STATISTICAL INDEPENDENCE

Another important and useful concept is that of STATISTICAL INDEPENDENCE. Two events are called statistically independent if knowledge of the occurrence of one does not affect the probability of occurrence of the other. What this means for events A and B, where $P(A) > 0$ and $P(B) > 0$, is that $P(B|A) = P(B)$ and $P(A|B) = P(A)$. From conditional probability it is known that

$$P(B|A) = \frac{P(B \text{ and } A)}{P(A)} \quad \text{or}$$

$$P(B \text{ and } A) = P(B \cdot A) = P(B|A) \cdot P(A),$$

but since $P(B|A) = P(B)$, it is said that A and B are <u>statistically</u> independent if

$$P(A \cdot B) = P(A) \cdot P(B) .$$

For example, if an air screen barrier and a destroyer barrier are placed about a convoy in such a manner that the probability of detection on one of the screens does not affect the probability of detection on the other screen, then the joint probability of detection of the submarine in <u>both</u> screens is simply the product of the probability of detection from each of the screens. Therefore, if the air screen barrier has a 0.50 probability of detection and the destroyer screen barrier has a 0.30 probability of detection, then the probability that the submarine will be detected on both screens is

P(Detection on <u>both</u> ASB and DSSB) = (0.50) (0.30) = 0.15.

This probability is, of course, contingent upon the submarine attempting to penetrate the screens. (The fact that this joint probability is smaller than either of the (marginal) probabilities which comprise it may, at first, seem intuitively incorrect; that is, it is easy to misinterpret the statement of what probability is being sought in such a problem. However, when it is realized that what is sought here is the probability that both platforms simultaneously hold contact on the submarine, the smaller resultant (joint) probability then seems more intuitively consistent.)

A major distinction between statistically independent events and mutually exclusive events is that the joint probability of the mutually exclusive events is zero, since the occurrence of one of a group of mutually exclusive events precludes the (simultaneous) occurrence of any other event in that group. The joint probability of independent events, on the other hand, is clearly not zero unless at least one of the independent events has zero probability of occurrence; If A and B are independent, then

$$P(A \text{ and } B) = P(A) \cdot P(B) = 0 \ ,$$

if and only if $P(A) = 0$ or $P(B) = 0$.

Supplementary solved problems covering elementary probability theory are presented at the end of Chapter 6 of Schaum's Outline Series, Theory and Problems of Statistics and at the end of Chapters 3 and 4 of Schaum's Outline Series, Theory and Problems of Probability. These problems should be consulted for a further review of the material presented in this chapter.

CHAPTER V: RANDOM VARIABLES

RANDOM VARIABLES

A variable which assumes a numerical value determined by the outcome of a probability experiment is called a random variable. Actually a random variable is a shorthand notation for listing each possible outcome of an experiment. For example, if a coin is tossed two times and the number of heads is counted, then the random variable may assume the value 0, 1, or 2.

There are two types of random variables, discrete and continuous. A discrete random variable is a random variable X on a sample space which has a finite or countably infinite number of outcomes; say, X(S) = $x_1, x_2, \ldots, x_n$. We can make X(S) into a probability space (conditions on this are listed with the probability distribution below) by defining the probability of x_i to be $P(x = i)$, which may also be written $f(x_i)$. This function f on X(s), defined by $f(x_i) = P(X = x_i)$, is called the distribution or probability function of X and is usually given in the form of a table, as shown in Table 5. (It should be noted that in the discrete case the distribution is the probability function developed in Chapter IV.)

x_1	x_2	$\cdots$	x_n
$f(x_1)$	$f(x_2)$	$\cdots$	$f(x_n)$

Table 5: Distribution Function of X

The distribution f satisfies the conditions:

$$f(x_i) \geq 0 \text{ and } \sum_{i=1}^{n} f(x_i) = 1$$

For example: suppose a pair of fair dice is tossed. We obtain the finite equiprobable space S consisting of the 36 ordered pairs of numbers between 1 and 6:

$$S = (1,1),\ (1,2),\ (1,3), \ldots,\ (6,6)$$

Let X assign to each point (a,b) in S the maximum of a and b; that is, X(a,b) = MAX (a,b) = X(S) = 1, 2, 3, 4, 5, 6. We compute the distribution f of X:

f(1) = P(X = 1) = P(1,1) = 1/36

f(2) = P(X = 2) = P((2,1), (2,2), (1,2)) = 3/36

f(3) = P(X = 3) = P((3,1), (3,2), (3,3), (2,3), (1,3)) = 5/36

and similarly f(4) = P(X = 4) = 7/36, f(5) = P(X = 5) = 9/36 and f(6) = P(X = 6) = 11/36. The information is put into the form of a table as shown in Table 6.

x_i =	1	2	3	4	5	6
$f(x_i)$ =	1/36	3/36	5/36	7/36	9/36	11/36

Table 6: Distribution f of X

It can also be seen that $f(x_i) \geq 0$ and

$$\sum_{i=1}^{6} f(x_i) = 1/36 + 3/36 + 5/36 + 9/36 + 11/36 = 36/36 = 1.$$

A <u>continuous</u> random variable is one in which the set of numbers X(S) is an interval or a set of intervals. If the interval $(a \leq X \leq b)$ is an event in S, then the probability would be $P(a \leq X \leq b)$ and is shown in Figure 15.

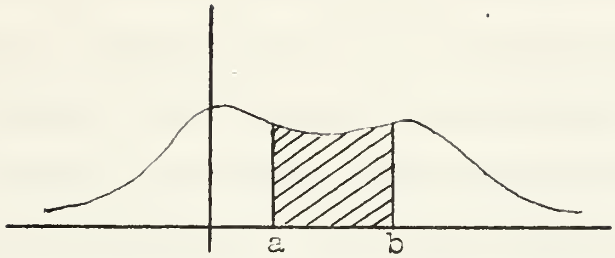


Figure 15: $P(a \leq X \leq b)$ = Area of Shaded Region

Where $P(a \leq x \leq b)$ corresponds to the area under the graph of f between $x = a$ and $x = b$. In the language of calculus, $P(a \leq x \leq b) = \int_a^b f(x)\,dx$, which in effect can be approximated by adding up areas under the curve between a and b. For example, take the area shown in Figure 15 and partition it into smaller areas as shown in Figure 16a and b, and then sum the areas of the resulting rectangles.

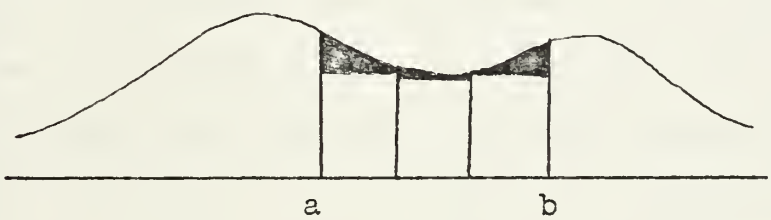


Figure 16a

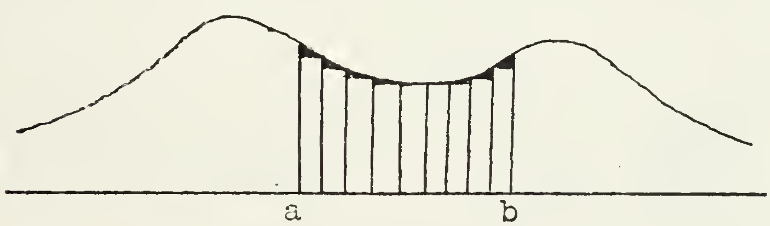


Figure 16b

An approximation of the area of each subregion of Figure 16a and 16b can now be made by taking the area of the rectangle which has the same base as the subregion and height equal to the minimum height for that subregion. By adding up the areas of the approximating rectangles, an approximation is now obtained for the area of the total region between a and b. The area that has not been included is the shaded area, which makes the results only an approximation. It can be seen that, due to the increased number of partitions each with a smaller base of Figure 16b, the shaded area is smaller, which results in a better approximation. Actually the greater the number of partitions, the smaller the shaded area, and the better the approximation. If f(x) is known, the exact area between a and b may easily be determined by using calculus. However, it should be noted that the use of calculus is not required for most practical cases of interest in this course, as values are obtained from standard tables. Standard tables exist for most continuous random variable distributions used in ASW applications; therefore, it is not intended that the student have a working knowledge of calculus. The following example is given, therefore, simply to show an appreciation of calculus:

Let X be a continuous random variable with the following distribution:

$$f(x) = \begin{cases} x/2 & \text{if } 0 \leq x \leq 2 \\ 0 & \text{elsewhere} \end{cases}$$

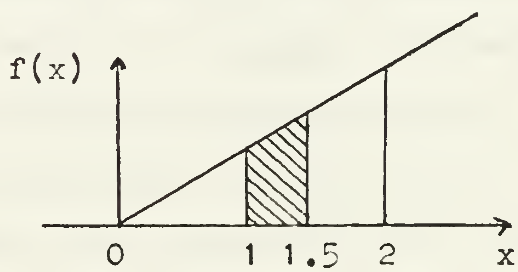


Figure 17: Graph of f

Then $P(1 \leq X \leq 1.5) = \int_a^b f(x)dx = \int_1^{1.5} \frac{1}{2}x dx = \frac{x^2}{4} \Big|_1^{1.5} = .5625 - .25 = .3125,$

where it is seen that the calculus operation of integration has been utilized to "sum" the area of the region under the curve between 1 and 1.5 as shown in Figure 17.

DENSITY FUNCTION

The function that describes the probability of each interval on the x-axis (for example, the bimodal function, f(x), in Figure 15) is called a continuous probability function or simply the (probability) density function of X. The density function must satisfy the constraints that $f(x) \geq 0$ for all x, and that the total area under the density function must equal 1. Special notice should be given the fact that the probability of a point-value of the random variable X is strictly equal to zero, even though the value of f at the point x may be non-zero. An intuitively appealing explanation of why the probability of a single outcome of a continuously distributed random variable is zero is given in the following example.

Consider an attack submarine (SSN) searching for enemy submarines. The probability that an initial detection bearing will be exactly some prescribed angle is zero. The key word here is exactly. For example, the word "exactly" would mean that a prescribed bearing of 30^o is exactly that right down to every decimal place. If the "exactly" is not used, then "30^o" may represent a whole event, which would have a positive probability. For example, if we are rounding off to the nearest degree and the density function is constant over the interval from 0^o to 360^o, then "30^o" would mean the event from $29.5000...^o$ to $30.499...^0$, which has

probability 1/360. If we are rounding off to the nearest minute, then "30°" refers to the interval from 29°59.500...' to 30°0.599...', which has probability 1/21,600. If we round off to the nearest second, then "30°" is the event from 29°59'59.500" to 30°0'0.499...", which has probability 1/1,296,000. "Exactly 30°" is a single point, not an interval, and its probability must be smaller than the probability of every interval that contains it, no matter how small. Thus the probability of a bearing angle of "exactly 30°" is zero.

CUMULATIVE DISTRIBUTION FUNCTION

Let X be a random variable, x a real number, and F(x) the probability that X takes on values less than or equal to x:

$$F(x) = P(X \leq x)$$

Then the function F is called the cumulative distribution function (CDF) of X. CDF values are of particular importance in ASW in their application to search regions. For example, (in the continuous random variable case) it is important to know what the probability of contact on a target along the search length, L, is up to and including the point X = L/2, where exactly half (or 1/3 or 3/4 or some other portion) of the area has been searched. This type of information is essential to allow reasonable allocation of available search effort. Further, the CDF of the target evaluated at the x_H value corresponding to that point half-way (in this example) through the search region represents this desired probability (cumulative), as seen in Figure 18, where, for purposes of generality, x is taken to be the middle or center x value of the search region's axis.

$F(X_H) = P[X \leq x_H]$

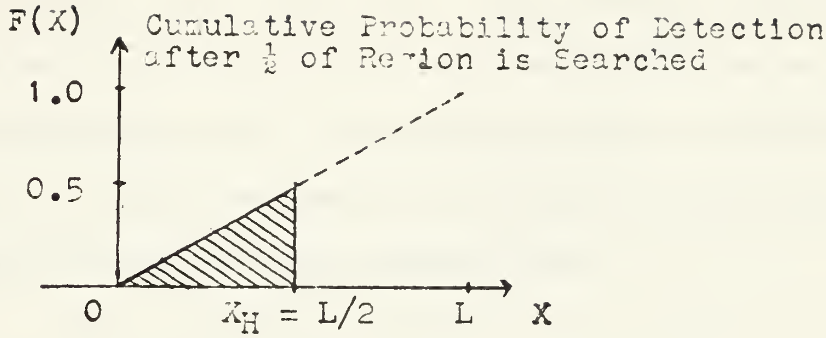


Figure 18: CDF of Search Region

For example, using the data shown in Table 6, the cumulative distribution function F of X would be as shown in Figure 19 which consists of summing the frequencies, $f(x_i)$, corresponding to those values of x less than or equal to x_i.

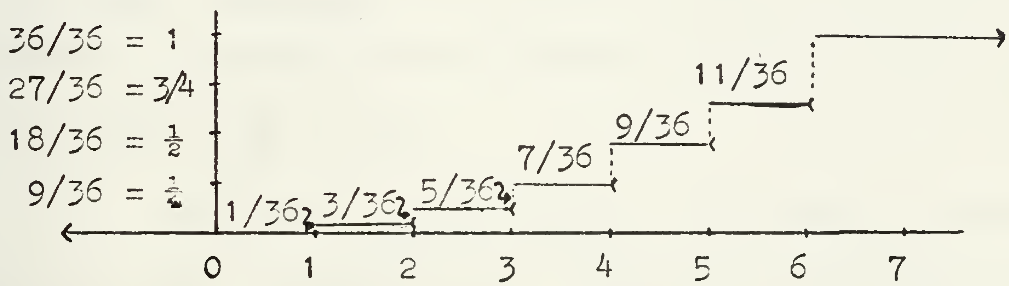


Figure 19: Graph of F

Observe that F in Figure 19 consists of a sequence of horizontal line segments and two rays. One ray coincides with that part of the x-axis to the left of x = 1 because $P(X \leq x)$ is 0 if $x < 1$. There is a jump in the graph at x = 1 because F(x) = 0 to the left of 1 and F(1) = 1/36. The graph has constant height F(x) = 1/36 for $1 \leq x \leq 2$. At x = 2 there is a jump of 3/36 (or f(2) = 3/36) bringing F(x) up to F(2) = 4/36. As we proceed along the graph from left to right, we find a jump at each integer x = 1, 2, 3, 4, 5, 6. For $x \geq 6$, the graph has another ray

extending indefinitely to the right because $F(x) = P(X \leq x)$ is 1 if X is any real number greater than or equal to 6. If we desire to find the CDF for a number between the integer values, for example F(5/2) we can see from Figure 19 that F(5/2) = 4/36. To compute the CDF of a continuous random variable, it is again necessary to use calculus; however, its values can usually be obtained from a standard table.

MEAN, VARIANCE AND STANDARD DEVIATION

The <u>mean</u>, <u>variance</u> and <u>standard deviation</u> are also used to describe a random variable. These will be shown here for discrete random variables only, as the values for continuous random variables will be found from tables.

The <u>mean</u> or <u>expectation</u> of X, denoted by E(X) (or simply E or μ) is defined, for discrete random variables, by $E(X) = x_1 f(x_1) + x_2 f(x_2) + \cdots + x_n f(x_n) = \sum_{i=1}^{n} x_i f(x_i)$.

That is, E(X) is the <u>weighted average</u> of the possible values of X, where each value is weighted by its probability. For example, to compute the mean of X for the data of Table 6,

$E(X) = \sum_{i=1}^{6} x_i f(x_i) = 1(1/36) + 2(3/36) + 3(5/36) + 4(7/36) + 5(9/36) + 6(11/36) = 161/36 = 4.47$.

The mean of a random variable X measures, in a certain sense, the "average" value of X.

The <u>variance</u> of X measures the "spread" or "dispersion" of X. Let X be a random variable with the distribution shown in Table 5. Then the <u>variance</u> of X, denoted by Var(X), is defined by $\mathrm{Var}(X) = \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i) = E((X - \mu)^2)$ where μ is the mean of X. For example, using the data

from Table 6 $Var(X) = \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i) = (1 - 4.47)^2\ 1/36 +$
$(2 - 4.47)^2\ 3/36 + (3 - 4.47)^2\ 5/36 + (4 - 4.47)^2\ 7/36 + (5 - 4.47)^2\ 9/36 +$
$(6 - 4.47)^2\ 11/36 = .334 + .508 + .3 + .043 + .07 + .715 = 1.97$.

Since the <u>standard deviation</u> of X, denoted by σ_x, is the (nonnegative) square root of Var(X): $\sigma_x = \sqrt{Var(X)}$. For example, using the data from Table 6, the standard deviation

$$\sigma_x = \sqrt{Var(X)} = \sqrt{1.97 =} 1.4 \ .$$

CHAPTER VI: BASIC STATISTICAL DISTRIBUTIONS

BACKGROUND

Generally, systems which arise in solving many practical ASW problems must be describable in terms of a simple mathematical model, and in most interesting cases this mathematical description is in probabilistic (rather than deterministic) terms. Before a problem can be analyzed for probabilistic solutions the basic unknown quantities must be represented as random variables (which constitute the mathematical model). Once this is done, the density function or the CDF is determined, and the system is thereby completely described in probabilistic terms. The only task remaining then is to compute the probabilities which constitute a problem solution, and this in itself is theoretically straight forward. However, it should be noted that some problems can be modeled more easily than others; yet an important advantage of this method is the possibility of describing many ASW problems in terms of a few "well known" random variables. The distributions which are developed here are the Bernoulli, Binomial, Chi-Square, Uniform, Normal, Bivariate Normal, Time-Varying Normal, Exponential and Poisson.

BERNOULLI DISTRIBUTION

A random variable X with exactly two possible values, say 1 and 0, where the probabilities $P\left[X = 1\right] = p$ and $P\left[X = 0\right] = 1 - p = q$, is called a Bernoulli random variable. Such a random variable is important because it may be used to describe any random experiment with two possible

outcomes, such as target detection from a single ping of an active sonar. Indeed, it is the Bernoulli distribution that is used in making a decision requiring a "Yes" or "No" answer. However, for our purposes Bernoulli random variables are used primarily as the basis for the Binomial distribution, which is discussed next.

BINOMIAL DISTRIBUTION

The Binomial distribution uses the two alternatives of the Bernoulli distribution, $P\left[X = 1\right] = p$ and $P\left[X = 0\right] = q$, and an n-fold repetition of the basic experiment. The problem now is to find the probability of exactly i successes in n repeated independent trials; for example, the probability of exactly two detections in six pings of an active sonar, or the probability of one detection in four sweeps of a radar. The Binomial random variable with parameters n and p is the discrete random variable given by $P\left[X = i\right] = \binom{n}{i} p^i q^{n-i}$, $i = 0, 1, 2, \ldots, n$, and $P\left[X = i\right] = 0$, otherwise. (Recall that $\binom{n}{i}$ was defined in Chapter III as $\frac{n!}{i!(n-i)!}$.

In order to use the Binomial distribution to find the probability of exactly 2 (or i) detections in 6 (or n) pings of an active sonar, the probability of a single ping success must be known. For example, if the probability of success is 0.6, or $p = 0.6$, then $q = 1 - p = 0.4$ and therefore $P\left[X = 2\right] = \binom{n}{i} p^i q^{n-i} = \binom{6}{2} (0.6)^2 (0.4)^{6-2} =$

$\frac{6!}{2!\,(6-2)!} (0.6)^2 (0.4)^4 = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 4 \cdot 3 \cdot 2 \cdot 1} (0.36)\,(0.0256) =$

$(15)\,(0.36)\,(0.0256) = 0.138$.

The mean of a Binomial random variable with parameters n and p is $\mu = np$, the variance is $\sigma^2 = npq$ and the standard deviation $\sigma = \sqrt{npq}$. In the above example, $\mu = np = (6)\,(0.6) = 3.6$, $\sigma^2 = npq = (6)\,(0.6)\,(0.4) = 1.44$ and $\sigma = \sqrt{1.44} = 1.2$. With a probability of 0.6 and 6 pings, we

would expect 3 to 4 detections from every 6 pings on the average; for this problem σ indicates, with a 68% confidence, that the number of detections from 6 pings is between 2.4 and 4.8 on the average.

The following problems show how the binomial distribution can be used;

1. A P-3C aircraft has been loaded with eight active sonobuoys which are known to have a 20% failure rate. These buoys are to be utilized in a situation requiring that failed buoys be replaced immediately to maintain pattern integrity. As this requires fast, coordinated action by the aircrew to reposition the aircraft for "delivery" and to prepare replacement buoys, it is important for them to estimate the number of buoys requiring replacement. Determine the probability that, out of four sonobuoys selected for a pattern,

(a) 1 sonobuoy will fail

(b) 0 sonobuoys will fail

(c) 2 sonobuoys will fail

(d) <u>at most</u> 2 sonobuoys will fail.

(Note that part (d) can be simplified through use of parts (a), (b), (c), and the probability complement concept.)

Answer: (a) $P\left[X = 1\right] = \binom{4}{1} (0.8)^3 (0.2)^1 = \frac{4!}{1!\ (4 - 1)!} (0.2)(0.8)^3$

$$= (4)\ (0.2)\ (0.51) = 0.4096$$

(b) 0.4096

(c) 0.1536

(d) $P\left[X \leq 2\right] = 1 - P\left[X > 2\right] =$

$$1 - \left[\binom{4}{3} (0.2)^3 (0.8)^1 + \binom{4}{4} (0.2)^4 (0.8)^0\right] = 0.9728$$

2. If the probability of damaging a submarine with a single depth charge from a destroyer is 0.5, what is the probability that 6 out of 10 depth charges will damage the submarine?

Answer: 0.21

3. Compute the (a) mean, (b) variance, and (c) standard deviation of a binomial distribution in which $p = 0.7$ and $n = 60$.

Answer: (a) 42

(b) 12.6

(c) 3.55

4. Consider a box of 200 fuses. It is known that for each fuse the probability of being defective is 0.02. What is the probability that at most six fuses are defective? (Set up the problem but do not carry out the calculations; although the solution is not too difficult if one has the use of a pocket calculator. It will be seen later that a good approximation to this answer can be obtained much easier by means of the Poisson approximation.)

Answer:

$$P\left[X = 6\right] = \sum_{i=0}^{6} \frac{n!}{i!(n-i)!} p^i q^{n-i}$$

$$= \sum_{i=0}^{6} \frac{200!}{i!(200-i)!} (0.02)^i (0.98)^{200-i}$$

$$= 0.8914$$

UNIFORM PROBABILITY DISTRIBUTION

The Uniform probability distribution represents situations in which any two sub-regions of the same size have equal probabilities, as explained below. This distribution is shown in Figure 20, which shows that the random variable X may assume only values between $X = a$ and $X = b$, or $a \leq X \leq b$, with non-zero probability (the probability of any region outside the interval $a \leq X \leq b$ is zero).

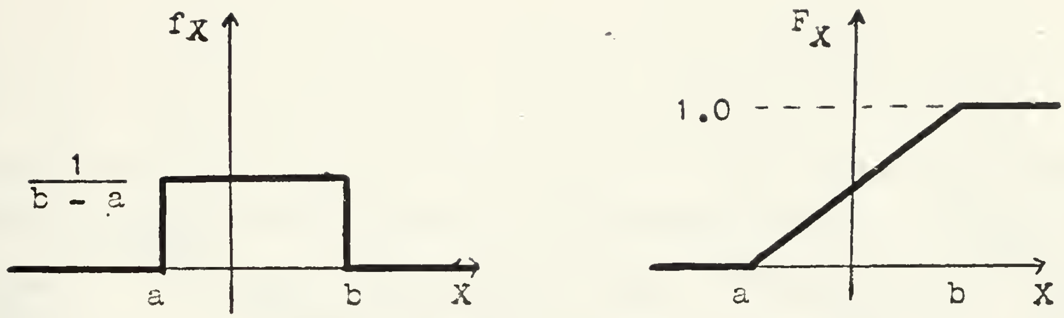


Figure 20: Uniform Probability Distribution and Its CDF

Since the area under a probability distribution must equal one, height of the density function between a and b must be $\frac{1}{b - a}$. The mean value of a Uniform distribution coincides with the central x-value, and is $(b + a)/2$. The variance is $(a - b)^2/12$. The figure also shows the cumulative distribution function for the Uniform distribution.

A Uniform distribution is described by writing "Uniform (a,b) or U(a,b) distribution"; for example, a Uniform distribution having its non-zero probability values in the interval 0 to 1 is denoted as "a Uniform zero, one distribution."

An example application of the Uniform random variable is the probability distribution of the location of a submarine in a narrow or restricted body of water. If it is known that the submarine will not be in water less than 100 feet deep, then the probability that the submarine will be in such regions is zero. If there is no information to indicate any preference in location, it must be assumed that the probability distribution for the location of the submarine is uniform, as shown in Figure 21.

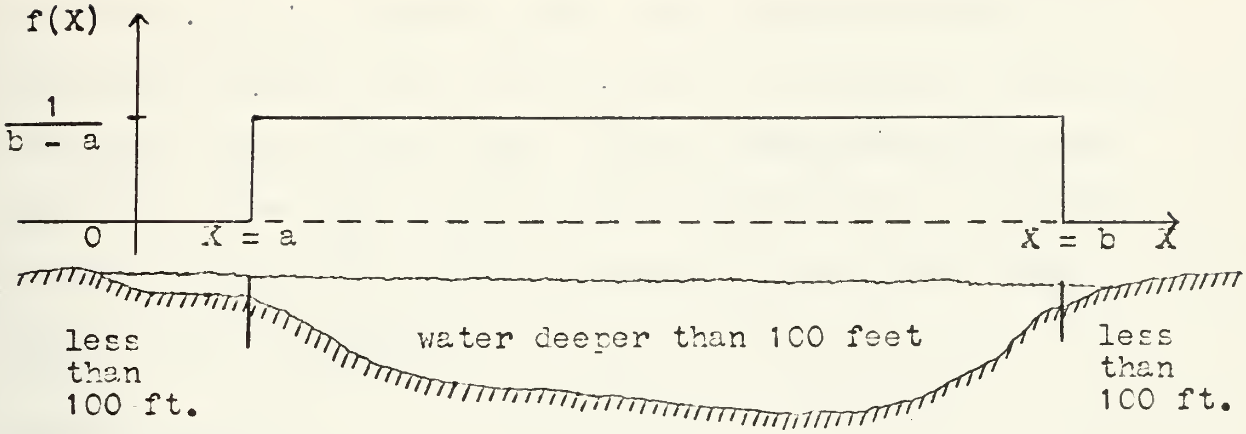


Figure 21: Profile of a Narrow Channel and Associated Uniform Probability Distribution

Suppose that a = 1,500 yards and b = 20,000 yards from a reference position (at one of the restriction boundaries, for example). The probability that the submarine is located in any subinterval, say of width 1,000 yards, is

$$
\begin{aligned}
P\left[c \leq X \leq d\right] &= (d - c)\,\frac{1}{(b - a)} \\
&= (1{,}000 \text{ yards})\,\frac{1}{18{,}500} \text{ yards} \\
&= 0.05
\end{aligned}
$$

Special note should be made of the fact that the dimensions (yards, in this example) associated with the physical problem must cancel so that the resultant probability is a dimensionless number. The probabilities of any two intervals of the same length within the interval from a to b are equal; that is,

$$P\left[c \leq X \leq d\right] = P\left[e \leq X \leq f\right] \text{ if } d - c = f - e,$$

where both $c \leq X \leq d$ and $e \leq X \leq f$ are within $a \leq X \leq b$.

Suppose that two sensors at separate locations each produce a line of bearing (LOB) to a submarine target. Each of the LOB's has a bearing error of ten degrees ($\pm$ 5°) associated with it. Suppose further that the

LOB's intersect in a plane region "centered" 117.5 n.m. from the first sensor and 66.5 n.m. from the second. The information available to a prosecuting platform is that the location of the submarine is Uniformly distributed over the plane region. Since a bearing error of 1 degree represents a range error of about 1 n.m. (perpendicular to the LOB) at a range of 60 n.m., the area is a trapezoid as shown by the "cross-hatched" region in Figure 22.

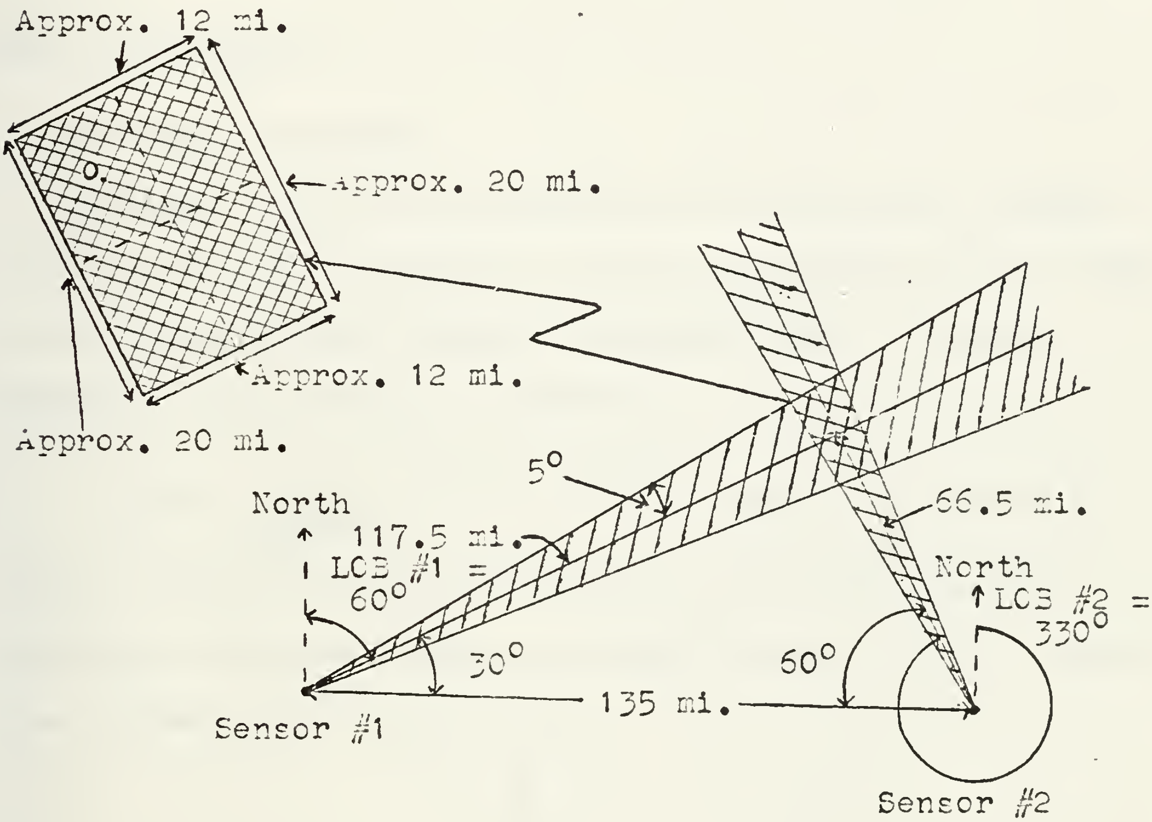


Figure 22: Trapezoid Produced by Two Sensors

Suppose sensors #1 and #2 are separated by 135 n.m. and the LOB's are 60 degrees from sensor #1 and 330 degrees from sensor #2; then the trapezoid <u>approximates</u> a rectangle with sides 20 by 12 n.m.; therefore, the probability region has an area of about 240 n.m.^{2}. If a sonobuoy with a 3 mile radius of detection is dropped in the probability region,

(a) what probability of detection does this sonobuoy yield?

(b) what probability of detection would six nonoverlapping sonobuoys yield?

Answer: (a) Probability of Detection = $\frac{\text{area of sonobuoy}}{\text{area of trapezoid}}$

$$\approx \frac{\pi r^2}{240} = \frac{28.27}{240} = 0.12$$

(b) Probability of Detection = 0.71

NORMAL DISTRIBUTIONS

THE NORMAL DISTRIBUTION

The Normal distribution is probably the most important continuous random variable used in ASW. It is used to describe the sum of a large number of statistically independent random variables. The Normal probability density function is defined as:

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x - \mu)^2/\sigma^2} \quad \text{for all real numbers } x,$$

where μ and σ^2 are the mean and variance. We denote the Normal distribution with mean μ and variance σ^2 by N(μ , σ^2). Figure 23 shows a Normal distribution of N(0,1).

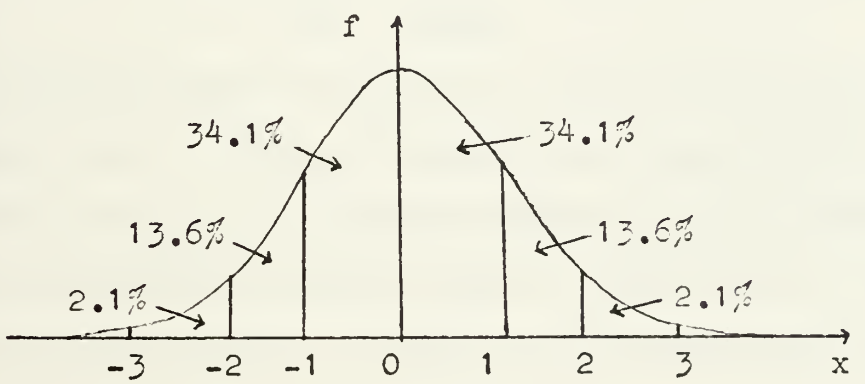


Figure 23: Normal Distribution N(0,1)

The probability that a Normally distributed random variable X is less than or equal to b, (which is the Cumulative Density Function) is given by:

$$F_X(b) = P\left[X \leq b\right] = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x - \mu)^2/\sigma^2} dx.$$

Evaluation of this intergal can be circumvented by use of standard tables. The probability that X is less than or equal to b can be represented by the shaded area in Figure 24, and its magnitude is determined with the aid of tables in Appendix A.

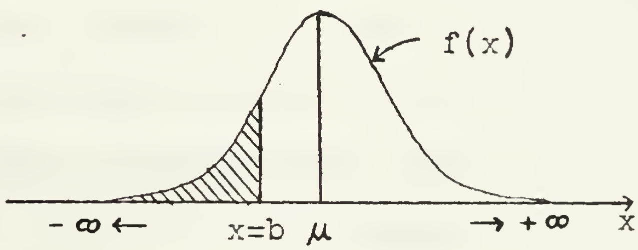


Figure 24: The Density Function of the Normal Distribution

Note that the Normal density function is dependent only on the two parameters μ and σ^2. If μ and σ^2 are both specified, probabilities can be obtained from the table in Appendix A. If

$$f_X(b) = \frac{1}{\sqrt{2\pi}} e^{-b^2/2} .$$

then the random variable X has a Normal distribution with mean 0 and variance 1, and X is called a Standard Normal random variable. Most tables are written for the Standard Normal Distribution; therefore, to obtain probabilities for any other mean or variance values the appropriate conversions must be made. To compute the probability that X lies between a and b, convert a and b into standard units:

$$a' = (a - \mu)/\sigma \qquad \text{and} \qquad b' = (b - \mu)/\sigma$$

respectively.

Then

$P(a \leq X \leq b) = P(a' \leq X^* \leq b')$ = area under the Standard Normal curve between a' and b', where X^* is the Standard Normal random variable.

The following problems are used to familiarize the student with the use of the Standard Normal distribution N(0,1) and Appendix A:

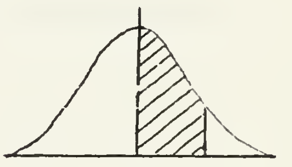
1. Let X be a random variable with the Standard Normal distribution N(0,1).

Find: (i) $P(0 \leq X \leq 1.42)$ (iv) $P(0.65 \leq X \leq 1.26)$

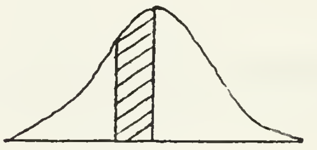
(ii) $P(-0.73 \leq X \leq 0)$ (v) $P(-1.79 \leq X \leq -0.54)$

(iii) $P(-1.37 \leq X \leq 2.01)$ (vi) $P(X \geq 1.13)$

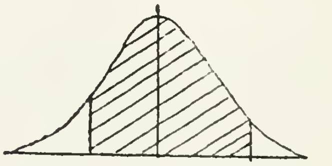
(i) $P(0 \leq X \leq 1.42)$ is equal to the area under the Standard Normal curve between 0 and 1.42. Thus in Appendix A, page A-1, look down the second column until 1.42 is reached. The entry is .4222. Hence $P(0 \leq X \leq 1.42)$ = .422.



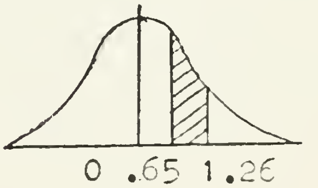
(ii) By symmetry, $P(-0.73 \leq X \leq 0) = P(0 \leq X \leq 0.73) = .2673$



(iii) $P(-1.37 \leq X \leq 2.01) = P(-1.37 \leq X \leq 0) + P(0 \leq X \leq 2.01) = .4147 + .4778 = .8925$



(iv) $P(0.65 \leq X \leq 1.26) = P(0 \leq X \leq 1.26) - P(0 \leq X \leq 0.65) = .3926 - .2422 = .1540$



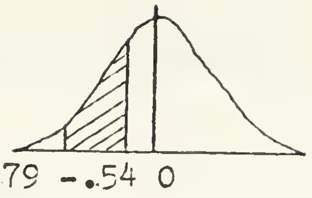
(v) $P(-1.79 \leq X \leq -0.54) =$

$P(0.54 \leq X \leq 1.79) =$

$P(0 \leq X \leq 1.79) -$

$P(0 \leq X \leq 0.54) =$

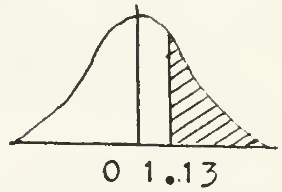
$.4633 - .2054 = .2579$



(vi) $P(X \geq 1.13) =$

$P(X \geq 0) - P(0 \leq X \leq 1.13) =$

$.5000 - .3708 = .1292$



2. Let X be Normally distributed with mean 8 and standard deviation 4. Find:

(i) $P(5 \leq X \leq 10)$ (iii) $P(X \geq 15)$

(ii) $P(10 \leq X \leq 15)$ (iv) $P(X \leq 5)$

(i) $P(5 \leq X \leq 10)$, must first convert to the standard normal distribution N(0,1)

$P(a \leq X \leq b) = P(a' \leq X^* \leq b')$

$a' = (a - \mu)/\sigma = (5 - 8)/4 = -.75$

$b' = (b - \mu)/\sigma = (10 - 8)/4 = .5$

therefore have $P(-.75 \leq X \leq .5) = P(-.75 \leq X \leq 0) + P(0 \leq X \leq 0.5) =$

$.2734 + .1915 = .4649$

(ii) $P(10 \leq X \leq 15) = P(.5 \leq X^* \leq 1.75) =$

$P(0 \leq X \leq 1.75) - P(0 \leq X^* \leq .5) =$

$.4599 - .1915 = .2684$

(iii) $P(X \geq 15) = P(X^* \geq 1.75) =$

$P(X^* \geq 0) - P(0 \leq X^* \leq 1.75) =$

$.5000 - .4599 = .0401$

(iv) $P(X \leq 5) = P(X^* \leq -.75) =$

$P(X^* \leq 0) - P(-.75 \leq X \leq 0) =$

$.5000 - .2734 = .2266$

3. Suppose the temperature T during June is Normally distributed with mean 68° and standard deviation 6°. Find the probability p that the temperature is between 70° and 80°.

70° in standard units = (70 - 68)/6 = .33

80° in standard units = (80 - 68)/6 = 2.00

Then $p = P(70 \leq T \leq 80) = P(0.33 \leq T \leq 2) =$

$P(0 \leq T^{*} \leq 2) - P(0 \leq T^{*} \leq 0.33) =$

$.4772 - .1293 = .3479$

Here T^{*} is the standardized random variable corresponding to T, so T^{*} has the Standard Normal distribution N(0,1).

4. Suppose that the lifetime of a certain type of electronic tube is a Normally distributed random variable. Suppose further that brand A has a lifetime X which is N(30 hours, 36 hours^2), while brand B has a lifetime Y which is N(45 hours, 9 hours^2). Of these two brands, which would you choose for use in an experimental aircraft with:

(a) 34 hours mission time

(b) 40 hours mission time

(Hint: you should be able to solve these questions by simply sketching the two Normal distributions).

BIVARIATE NORMAL DISTRIBUTION

The Bivariate Normal and the Time-Varying Normal distributions are proposed inclusions to the computer program of the P-3 aircraft for tactical, on-station use of probability theory, and although a thorough development of the distributions is considered beyond the scope of this course, the following cursory description of each of these distributions should be sufficient for an intuitive idea of how these distributions are used.

The Bivariate Normal distribution is the joint probability distribution of two Gaussian (or Normal) random variables. Basically, this distribution is the three-dimensional extension of the two-dimensional (the value and its frequency) Normal distribution. As seen in Figure 25, the Bivariate Normal distribution resembles a hat, and the <u>specific</u> Bivariate Normal distribution shown has a central value, which may or may not be located at the origin.

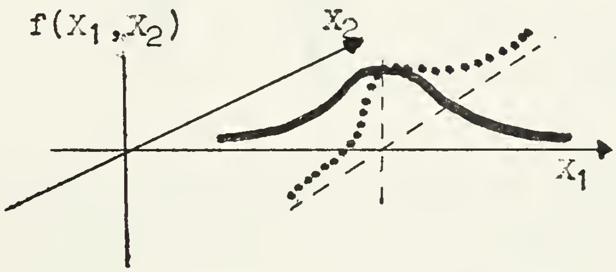


Figure 25: Bivariate Normal Distribution

The Bivariate Normal distribution is useful in Search Theory, and for two-dimensional area coverage.

If the student is interested in further development of the Bivariate Normal, reference should be made to <u>Selected Methods and Models in Military Operations Research</u>, P. W. Zehna, <u>et al</u>, U. S. G. P.O., 1971, pages 49 - 61.

TIME-VARYING NORMAL DISTRIBUTION

The Time-Varying Normal distribution describes a Normal random variable whose distribution varies as a function of time as well as of another parameter such as range. This distribution varies with time, and Figure 26 shows an example where the mean is zero for all values of time t, but the variance is increasing with increasing t.

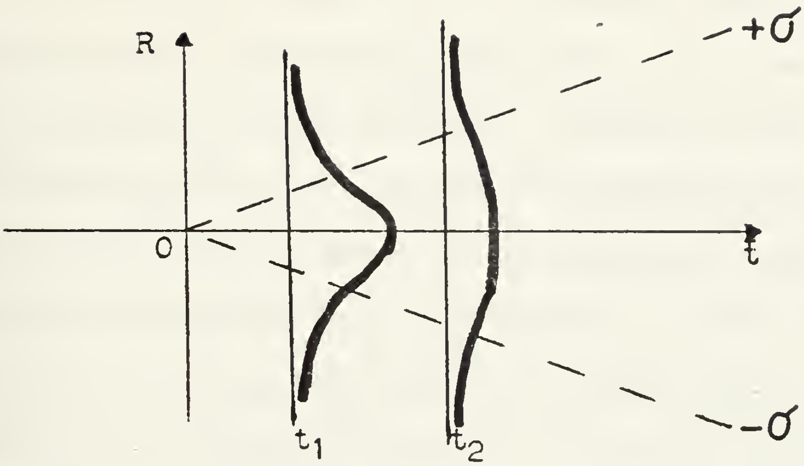


Figure 26: Time-Varying Normal Distribution

The distribution shown in Figure 26 can be used to describe a submarine opening a reported datum, where a "datum outbound" situation is assumed to exist. As the figure illustrates, the probability that the submarine is in close proximity to the reported datum--as reported at time t_0 or 0--decreases with increasing time. This is indicated by the spreading out and flattening of the distribution from times t_1 to t_2.

It is important to understand the development thus far presented to provide correct insight into what this distribution represents. Further, it is a definite advantage--and, in fact, a necessity for tactical, real-time usage--to have a digital computer to perform the mathematics associated with the Time-Varying Normal distribution resulting from an operator-input assessment of the tactical situation. The computer could then output the pertinent probabilities and options for operator use. Again, if the operator is to utilize the computer to employ the distribution, then it is imperative that he understand the tactical assumptions it is based on.

CHI-SQUARE DISTRIBUTION

The Chi-Square random variable is useful for ASW applications in that it can be used to describe <u>any</u> system that is the sum of the squares of Normal independent random variables. Probably the most important of these applications is that of describing <u>radial distance</u>, where the distance is in terms of two or three directions, <u>each</u> of which has distance distributed Normally. Therefore, the radial distance to a specific point (from the origin) is the <u>square root</u> of the sum of the squares of the cartesian coordinates of the point (by the Pythagorean theorem).

An example which displays the utility of the Chi-Square distribution is to obtain the probability that a submarine is within 5 miles of a central datum (the origin, for convenience), where it is known that the probabilities of the submarine being at a distance in miles along each of the x and y axis are each distributed N(0,100) due to error in the system that initially located the submarine. The submarine's position can be <u>in error</u> (or vary from the stated datum position) in <u>both</u> the x and y directions; therefore, its position may be specified by two random variables, X and Y. Although there are two random variables involved, the effect of using the Chi-Square distribution is a reduction of the problem to a single random variable R, which represents radial distance from the submarine to the central datum:

$$R = \sqrt{X^2 + Y^2} \ .$$

The problem is to find the probability of the shaded region shown in Figure 27.

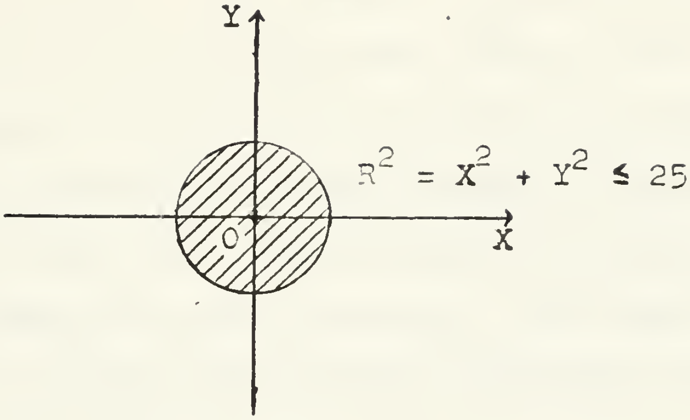


Figure 27: Sum of Squares Less Than or Equal to 25 Miles

That is, to calculate $P\left[(X^2 + Y^2)^{\frac{1}{2}} \leq 5\right]$, or equivalently, $P\left[X^2 + Y^2 \leq 25\right]$, we apply the following principle from probability theory (refer to Central Limit Theorem on page 142 of Schaum's Outline Series, Theory and Problems of Statistics): The sum of n independent Normally distributed random variables forms a new random variable known as the Chi-Square random variable with "n degree of freedom"; this is denoted by

$$\sum_{i=1}^{n} X_i^2 = \chi_n^2 .$$

In the specific two-dimensional case treated here, if we put

$$X_1 = \frac{1}{10} X \text{ and } X_2 = \frac{1}{10} Y, \text{ then } \sum_{i=1}^{2} X_i^2 = \chi_2^2.$$

The probability is now denoted by

$$P\left[X^2 + Y^2 \leq 25\right] = P\left[10^2 X_1^2 + 10^2 X_2^2 \leq 25\right] = P\left[X_1^2 + X_2^2 \leq \frac{25}{10^2}\right] =$$

$$P\left[X_1^2 + X_2^2 \leq 0.25\right], \text{ or } = P\left[\chi_2^2 \leq 0.25\right] = F_{\chi_2^2}(0.25).$$

The last term is the probability that the Chi-Square random variable with two degrees of freedom, is less than or equal to 0.25. This cumulative value is found from the "Percentage Points of the Chi-Square

Distribution" in Appendix A. This value, $F_{\chi_2^2}(0.25)$ is found by entering the table with n (the left-hand column) value of 2 (the number of degrees of freedom-- ν is used for n in the table), and reading directly to the right until the column is reached that has the tabulated value <u>closest to</u> 0.25. That column is then followed vertically upward until the top row is reached--the row is labled as Q. The value in this row and column is the complement of the desired probability, and it is

$$Q_{\chi_2^2}(0.25) = 0.90$$

$$1 - Q_{\chi_2^2}(0.25) = 1 - 0.90; \text{ therefore}$$

$$F_{\chi_2^2}(0.25) = 0.10.$$

Now, consider the problem of finding the probability that a submarine will be "killed" by a moored mine if it is known that the kill capability of the mine is a function of the radial distance between the mine and the submarine. This is a three dimensional problem, as the submarine depth is also considered. The kill function is known to be certain if the submarine is within 30 yards of the mine, and the probabilities that the submarine will pass within a given range in each of the three coordinate directions (depth, "North-South," and "East-West") are independently distributed N(0,625) in yards. The problem is to calculate $P\left[(X^2 + Y^2 + Z^2) \leq 900\right]$.

$$\text{Let } X_1 = \frac{X}{\sigma_X} = \frac{X}{25}$$

$$X_2 = \frac{Y}{\sigma_Y} = \frac{Y}{25}$$

$$X_3 = \frac{Z}{\sigma_Z} = \frac{Z}{25}$$

(This technique is that used to normalize and standardize a Normal random variable), then

$$\sum_{i=1}^{3} X_i^2 = \chi_3^2 .$$

So that,

$$P\left[X^2 + Y^2 + Z^2 \leq 30\right] = P\left[25^2(X_1^2 + X_2^2 + X_3^2) \leq 900\right] =$$

$$P\left[X_1^2 + X_2^2 + X_3^2 \leq \frac{900}{25^2}\right] = P\left[X_1^2 + X_2^2 + X_3^2 \leq 1.44\right] =$$

$$F_{\chi_3^2}(1.44) .$$

From the table on page A-3,

$$Q_{\chi_3^2}(1.44) = 0.75, \text{ and}$$

$$F_{\chi_3^2}(1.44) = 0.25 .$$

Hence, under the assumed probability distributions, the probability that the submarine will be "killed" is 0.25.

There are other methods of solving this problem which involve working with the random variables X and Y; however, the Chi-Square method is easier, and it is convenient because the probabilities associated with it can be simply obtained from a precomputed table.

The following problems will help familiarize the student with the use of the Chi-Square distribution and Appendix A:

1. Find $P\left[\chi_{20}^2 \leq 21.5\right] = F_{\chi_{20}^2}(21.5)$

   Answer: 0.75

2. Find $P\left[\chi_{20}^2 \leq 14.7\right]$

   Answer: 0.25

3. Find the value of X for which $P\left[\chi_{25}^2 \leq x\right] = 0.75$

   Answer: 29.3

4. An antiaircraft shell designed to burst at a specified point in space has x, y, and z errors in a coordinate system with origin at the specified point which are independent Normally distributed random variables X, Y, Z, each with mean 0 and standard deviation 50, with length measured in feet. Find the probability that the radial error in the burst point is at most 100 feet.

Answer: 0.75

EXPONENTIAL DISTRIBUTION

An exponential function is a function of the form $f(t) = e^t$, where e is a constant which is very much like π , except that e is approximately 2.7183, vice 3.1416. For example, if t = 2, then $f(2) = e^2 = (2.7183)^2 = 7.3891$, as seen from the last entry opposite x = 2.00 in the table of Appendix A, page A-2. The specific exponential function of interest here is the <u>negative exponential</u> or $f(t) = e^{-t}$ (see Appendix A, page A-2 for values of e^{-t}) which is the reciprocal of the positive exponential shown before, or $f(t) = 1/e^t$.

The Exponential distribution itself is actually $f_X(t) = \lambda e^{-\lambda t}$, (t= 0) where λ , the lower case Greek letter "lambda," is a positive constant. It turns out that $1/\lambda$ is the mean of this distribution. The graph of the Exponential distribution is shown in Figure 28. It is seen in the figure that the function (the Exponential distribution) "decays" with the progression of time.

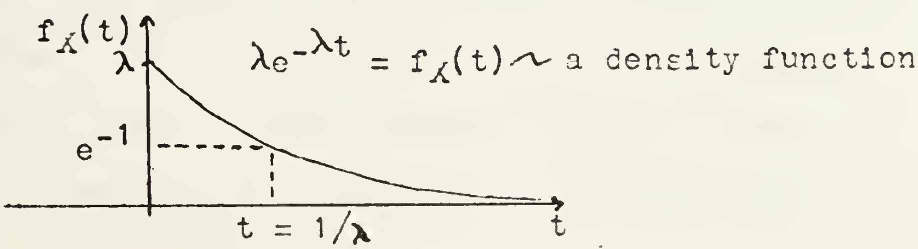


Figure 28: The Exponential Distribution

The Exponential distribution can be used to describe many phenomena connected with military applications. An example where such a function occurs is the MTTF (Mean Time To Failure) of an electronic equipment. The time axis represents the total lifetime of the equipment itself. The random variable X represents the total lifetime; that is, to say that X assumes a value t_0 means that the equipment fails at time t_0.

The cumulative distribution function is

$$F_X(t) = 1 - e^{-\lambda t}, \; t \geq 0$$

Therefore, the table in Appendix A, page A-2, is still usable, except that the value found in the table must be subtracted from one. Figure 29 shows the Cumulative Exponential probability distribution.

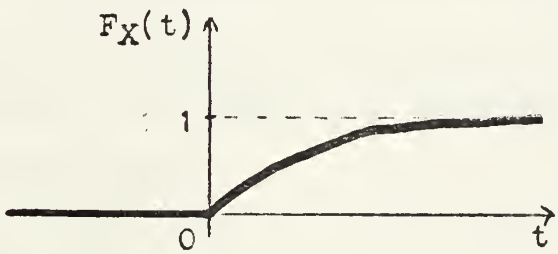


Figure 29: The Cumulative Exponential Distribution Function

Practical ASW applications of the Exponential probability distribution are predicting: (1) the <u>mean time to trail</u> a submarine and (2) the <u>mean time to fail</u> for ASW equipment.

The following problems show some of these applications.

1. Suppose that it is desired to trail a submarine that transits a barrier. Take the time that the submarine enters the barrier to be $t = 0$, and let P_0 be the probability that trail is initiated at $t = 0$, where P_0 is known. In this problem λ is the reciprocal of the mean time to lose trail (or the average rate of losing trail). For this

problem let $1/\lambda$ be 20 days, and let $P_0 = 0.5$. What is the probability the submarine is trailed at time t = 10 days? Since $1/\lambda = 20$,

$\lambda = 1/20 = 0.05$; therefore,

$$P(t) = P_0 (1 - e^{-\lambda t}) = 0.5 (1 - e^{-(0.5)(10)}) =$$

$$0.5 (1 - 0.6) = (0.5)(0.4) = 0.2$$

2. A device is to operate until failure. The pdf in a model describing the time of failure is $f(t) = e^{-t}, 0 < t$, where t is in hours. What is the probability the device will fail in the first two hours?

Answer: In order that e^{-t} be of the form $\lambda e^{-\lambda t}$, λ must equal one, that is, the cdf is $F(t) = 1 - e^{-t}$. Thus $P\left[\text{Fail in first 2 hrs}\right] =$ $1 - e^{-2} = 1 - 0.135 = 0.865$

3. A resistor is installed in a fire control system. The pdf describing the time to failure of the resistor is $f(t) = \lambda e^{-\lambda t}, 0 < t$ where t is in years. If the MTTF is 15 years what is the probability the resistor will last 6 years?

$$\lambda = \frac{1}{\text{MTTF}} = \frac{1}{15} = 0.07$$

$$P\left[\text{Resistor last 6 years}\right] = 1 - P\left[\text{Resistor fails in 6 years}\right] =$$

$$1 - \left[1 - \lambda e^{-\lambda t}\right] = 1 - \left[1 - 0.07 e^{-(0.07)(6)}\right] =$$

$$1 - \left[1 - (0.67)(0.3)\right] = 1 - 0.8 = 0.2$$

4. The reliability of a device during a mission depends on T, the time the device has been in storage before the start of the operation. In a probability model, A is the event the device survives the mission, $f(t) = e^{-t}$ is the pdf of T. Find P(A) for a 6 month mission, where T is in years.

$$F(t) = 1 - e^{-t}, \; P(A) = 0.6$$

POISSON DISTRIBUTION

Many physical phenomena can be described by the Poisson distribution. Emission of particles from a radiating source, arrivals of messages in a CIC, and submarines transiting a barrier are a few of the shipboard ASW situations that may require this distribution for analysis of the problem.

The basic notion of this distribution involves a stream of "customers" arriving at a "service point" at a constant average arrival rate. It is assumed also that arrivals are independent in the sense that the number of arrivals in any given time interval is independent of the number of arrivals in any other interval which does not overlap the first.

The Poisson distribution is given by

$$p(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots$$

where $\lambda > 0$ is some constant.

In the applications of the Poisson distribution that are of primary interest in ASW, the Poisson distribution represents the number of "arrivals" in the given time interval. For example, if time is measured in hours, the given time interval is a certain ten-hour period, and $\lambda = 0.5$, then the notation

$$p(2,0.5) = \frac{(0.5)^2 e^{-0.5}}{2!} = 0.076$$

is interpreted to mean:

"the probability of exactly two arrivals in the given time period is 0.076."

The length of the time interval is not apparent in the expression defining $p(k, \lambda)$. Suppose the basic unit of time (this may be a second,

a minute, an hour, a day, a year, etc.) has been chosen. Let T be the length of the given time interval, i.e. the number of basic time units. Then the λ in the Poisson distribution has the property that λ/T is the mean arrival rate, expressed as the number of arrivals per unit of time. That is, if we denote the mean arrival rate by α , then

$$\alpha = \lambda/T, \qquad \text{or } \lambda = \alpha T.$$

Therefore, for applications other than the Binomial approximation indicated next, we use

$$p(k, \alpha T) = \frac{e^{-\alpha T}(\alpha T)^k}{k!}, \quad k = 0, 1, 2, \ldots$$

to express the probabilities for the various possible number of arrivals.

This distribution has a mean, $\mu = \lambda$, variance, $\sigma^2 = \lambda$ and standard deviation, $\sigma = \sqrt{\lambda}$.

In addition to having its own applications, the Poisson distribution also provides us with a close approximation of the Binomial distribution for small k, provided that p is small and $\lambda = np$. This is indicated in Table 7 where n = 100, p = 1/100 and $\lambda = np = 1$.

k	0	1	2	3	4	5
Binomial	0.366	0.360	0.185	0.0610	0.0149	0.0029
Poisson	0.368	0.368	0.184	0.0613	0.0153	0.00307

Table 7: Comparison of Binomial and Poisson Values

The Poisson distribution can probably best be understood through examples and problems.

1. Find: (i) $e^{-1.3}$, (ii) $e^{-2.5}$.

From Appendix A, page A-2,

(i) $e^{-1.3} = 0.273$

(ii) $e^{-2.5} = 0.0821$

2. For the Poisson distribution $p(k, \lambda) = \dfrac{\lambda^k e^{-\lambda}}{k!}$, find

(i) $p(2,1)$, (ii) $p(3,\frac{1}{2})$, (iii) $p(2,.7)$

(i) $p(2,1) = \dfrac{1^2 e^{-1}}{2!} = \dfrac{e^{-1}}{2} = \dfrac{0.368}{2} = 0.184$

(ii) $p(3,\frac{1}{2}) = 0.013$

(iii) $p(2,.7) = 0.12$

3. Suppose that in a particular region submarine detection times obey a Poisson distribution, with a mean detection rate of one submarine every 12 days. Calculate the probability that five submarines will be detected in the operating area during a quarter (120 days)?

As indicated above, $\lambda = \alpha T$, where α represents the average detection rate (i.e. the average number of detections per day, which is 1/12), and T the length of the given time interval, so that T = 120). Therefore,

$$P\left[K = 5\right] = \frac{e^{-\alpha T}(\alpha T)^k}{k!} = \frac{e^{-(1/12)(120)}\,(1/12)\,(120)^5}{5!} =$$

0.03783

4. Using the information from problem 3, find the probability of detecting <u>five or fewer</u> (at most five) submarines during the quarter.

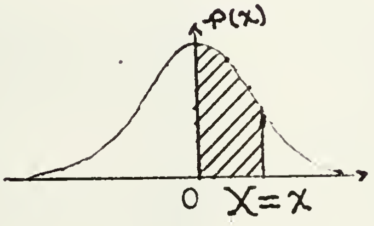
$$P\left[K = 5\right] = \sum_{k=0}^{5} \frac{\left[(1/12)\,(120)\right]^k e^{-(1/12)\,(120)}}{k!} = 0.6709$$

5. Approximate the Binomial distribution problem number 4 in the Binomial distribution section of this chapter, using the Poisson distribution. Answer: 0.9

APPENDIX A

Areas Under the Standard Normal Curve from 0 to .

x	P(x)	x	P(x)
0.00	0.0000	1.26	0.3962
0.33	0.1293	1.37	0.4147
0.50	0.1915	1.42	0.4222
0.54	0.2054	1.75	0.4599
0.65	0.2422	1.79	0.4633
0.73	0.2673	2.00	0.4772
1.13	0.3708		



EXPONENTIAL FUNCTION

x	e^x	e^{-x}
0.0	1.00000 00000 00000	1.00000 00000 00000 00000
0.1	1.10517 09180 75648	0.90483 74180 35959 57316
0.2	1.22140 27581 60170	0.81873 07530 77981 85867
0.3	1.34985 88075 76003	0.74081 82206 81717 86607
0.4	1.49182 46976 41270	0.67032 00460 35639 30074
0.5	1.64872 12707 00128	0.60653 06597 12633 42360
0.6	1.82211 88003 90509	0.54881 16360 94026 43263
0.7	2.01375 27074 70477	0.49658 53037 91409 51470
0.8	2.22554 09284 92468	0.44932 89641 17221 59143
0.9	2.45960 31111 56950	0.40656 96597 40599 11188
1.0	2.71828 18284 59045	0.36787 94411 71442 32160
1.1	3.00416 60239 46433	0.33287 10836 98079 55329
1.2	3.32011 69227 36547	0.30119 42119 12202 09664
1.3	3.66929 66676 19244	0.27253 17930 34012 60312
1.4	4.05519 99668 44675	0.24659 69639 41606 47694
1.5	4.48168 90703 38065	0.22313 01601 48429 82893
1.6	4.95303 24243 95115	0.20189 65179 94655 40849
1.7	5.47394 73917 27200	0.18268 35240 52734 65022
1.8	6.04964 74644 12946	0.16529 88882 21586 53830
1.9	6.68589 44422 79269	0.14956 86192 22635 05264
2.0	7.38905 60989 30650	0.13533 52832 36612 69189
2.1	8.16616 99125 67650	0.12245 64282 52981 91022
2.2	9.02501 34994 34121	0.11080 31583 62333 88333
2.3	9.97418 24548 14721	0.10025 88437 22803 73373
2.4	11.02317 63806 41602	0.09071 79532 89412 50338
2.5	12.18249 39607 03473	0.08208 49986 23898 79517
2.6	13.46373 80350 01690	0.07427 35782 14333 88043
2.7	14.87973 17248 72834	0.06720 55127 39749 76513
2.8	16.44464 67710 97050	0.06081 00626 25217 96500
2.9	18.17414 53694 43061	0.05502 32200 56407 22903
3.0	20.08553 69231 87668	0.04978 70683 67863 94298
3.1	22.19795 12814 41633	0.04504 92023 93557 80607
3.2	24.53253 01971 09349	0.04076 22039 78366 21517
3.3	27.11263 89206 57887	0.03688 31674 01240 00545
3.4	29.96410 00473 97013	0.03337 32699 60326 07948
3.5	33.11545 19586 92314	0.03019 73834 22318 50074
3.6	36.59823 44436 77988	0.02732 37224 47292 56080
3.7	40.44730 43600 67391	0.02472 35264 70339 39120
3.8	44.70118 44933 00823	0.02237 07718 56165 59578
3.9	49.40244 91055 30174	0.02024 19114 45804 38847
4.0	54.59815 00331 44239	0.01831 56388 88734 18029
4.1	60.34028 75973 61969	0.01657 26754 01761 24754
4.2	66.68633 10409 25142	0.01499 55768 20477 70621
4.3	73.69979 36995 95797	0.01356 85590 12200 93176
4.4	81.45086 86649 68117	0.01227 73399 03068 44118
4.5	90.01713 13005 21814	0.01110 89965 38242 30650
4.6	99.48431 56419 33809	0.01005 18357 44633 58154
4.7	109.94717 24521 23499	0.00909 52771 01695 81709
4.8	121.51041 75187 34881	0.00822 97470 49020 02884
4.9	134.28977 96849 35485	0.00744 65830 70924 34052
5.0	148.41315 91025 76603	0.00673 79469 99085 46710

PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION—VALUES OF χ^2 IN TERMS OF Q AND ν

ν \ Q	0.995	0.99	0.975	0.95	0.9	0.75	0.5	0.25
1	(-5) 3.92704	(-4) 1.57088	(-4) 9.82069	(-3) 3.93214	0.0157908	0.101531	0.454937	1.32330
2	(-2) 1.00251	(-2) 2.01007	(-2) 5.06356	0.102587	0.210720	0.575364	1.38629	2.77259
3	(-2) 7.17212	0.114832	0.215795	0.351846	0.584375	1.212534	2.36597	4.10835
4	0.206990	0.297110	0.484419	0.710721	1.063623	1.92255	3.35670	5.38527
5	0.411740	0.554300	0.831211	1.145476	1.61031	2.67460	4.35146	6.62568
6	0.675727	0.872085	1.237347	1.63539	2.20413	3.45460	5.34812	7.84080
7	0.989265	1.239043	1.68987	2.16735	2.83311	4.25485	6.34581	9.03715
8	1.344419	1.646482	2.17973	2.73264	3.48954	5.07064	7.34412	10.2188
9	1.734926	2.087912	2.70039	3.32511	4.16816	5.89883	8.34283	11.3887
10	2.15585	2.55821	3.24697	3.94030	4.86518	6.73720	9.34182	12.5489
11	2.60321	3.05347	3.81575	4.57481	5.57779	7.58412	10.3410	13.7007
12	3.07382	3.57056	4.40379	5.22603	6.30380	8.43842	11.3403	14.8454
13	3.56503	4.10691	5.00874	5.89186	7.04150	9.29906	12.3398	15.9839
14	4.07468	4.66043	5.62872	6.57063	7.78953	10.1653	13.3393	17.1170
15	4.60094	5.22935	6.26214	7.26094	8.54675	11.0365	14.3389	18.2451
16	5.14224	5.81221	6.90766	7.96164	9.31223	11.9122	15.3385	19.3688
17	5.69724	6.40776	7.56418	8.67176	10.0852	12.7919	16.3381	20.4887
18	6.26481	7.01491	8.23075	9.39046	10.8649	13.6753	17.3379	21.6049
19	6.84398	7.63273	8.90655	10.1170	11.6509	14.5620	18.3376	22.7178
20	7.43386	8.26040	9.59083	10.8508	12.4426	15.4518	19.3374	23.8277
21	8.03366	8.89720	10.28293	11.5913	13.2396	16.3444	20.3372	24.9348
22	8.64272	9.54249	10.9823	12.3380	14.0415	17.2396	21.3370	26.0393
23	9.26042	10.19567	11.6885	13.0905	14.8479	18.1373	22.3369	27.1413
24	9.88623	10.8564	12.4011	13.8484	15.6587	19.0372	23.3367	28.2412
25	10.5197	11.5240	13.1197	14.6114	16.4734	19.9393	24.3366	29.3389
26	11.1603	12.1981	13.8439	15.3791	17.2919	20.8434	25.3364	30.4345
27	11.8076	12.8786	14.5733	16.1513	18.1138	21.7494	26.3363	31.5284
28	12.4613	13.5648	15.3079	16.9279	18.9392	22.6572	27.3363	32.6205
29	13.1211	14.2565	16.0471	17.7083	19.7677	23.5666	28.3362	33.7109
30	13.7867	14.9535	16.7908	18.4926	20.5992	24.4776	29.3360	34.7998
40	20.7065	22.1643	24.4331	26.5093	29.0505	33.6603	39.3354	45.6160
50	27.9907	29.7067	32.3574	34.7642	37.6886	42.9421	49.3349	56.3336
60	35.5346	37.4848	40.4817	43.1879	46.4589	52.2938	59.3347	66.9814
70	43.2752	45.4418	48.7576	51.7393	55.3290	61.6983	69.3344	77.5766
80	51.1720	53.5400	57.1532	60.3915	64.2778	71.1445	79.3343	88.1303
90	59.1963	61.7541	65.6466	69.1260	73.2912	80.6247	89.3342	98.6499
100	67.3276	70.0648	74.2219	77.9295	82.3581	90.1332	99.3341	109.141
x	-2.5758	-2.3263	-1.9600	-1.6449	-1.2816	-0.6745	0.0000	0.6745

$$Q(\chi^2|\nu)=\left[2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)\right]^{-1}\int_{\chi^2}^{\infty} e^{-\frac{t}{2}}\,t^{\frac{\nu}{2}-1}\,dt$$

BIBLIOGRAPHY

Lipschutz, S., "Schaum's Outline of Theory and Problems of Probability," McGraw-Hill Book Company, 1965.

Spiegil, M. R., "Schaum's Outline of Theory and Problems of Statistics," McGraw-Hill Book Company, 1961.

Zehna, P. W., et al, "Selected Methods and Models in Military Operations Research," U. S. G. P.O., 1971, pages 49 - 61.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Chairman ASW Academic Group, Code 71 Naval Postgraduate School Monterey, California 93940	2
4.	Assoc. Professor Carroll O. Wilde, Code 28 Department of Mathematics Naval Postgraduate School Monterey, California 93940	1
5.	LT Billy Joe Smith, USN 102 Deakin Circle Monterey, California 93940	1
6.	LT Larry Wayne Vice, USN 1039 Halsey Drive Monterey, California 93940	1

160076
Thesis
S5737 Smith
c.1 Probability and its applications to anti-submarine warfare.